UNRAVELLING A VINE: A TAXONOMIC AND CONSERVATION GENETICS STUDY OF *TETRASTIGMA LOHERI* GAGNEP. (VITACEAE) IN THE PHILIPPINES

A thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Plant Biology

by Jasper John Aragon Obico

University of Canterbury

2019

CONTENTS

ACKNOWLEDGEMENTS

To Pieter Pelser, my supervisor, for entrusting me this PhD project and patiently mentoring me throughout my PhD journey. To Marie Hale and Tammy Steeves, my former and present co-supervisors respectively, for helping in the conservation genetics aspect of my study. To Julie Barcelona, my associate supervisor, for great advice and coordination of fieldwork in the Philippines. To the University of the Philippines-Manila for allowing me to leave work and study overseas. To the Ministry of Foreign Affairs and Trade of the New Zealand government for awarding me the New Zealand ASEAN PhD Scholarship. To the Rufford Foundation for partially funding my fieldwork in the Philippines. To the Koiata Botanical Trust for partially funding my laboratory work. To SBS technical and administrative staff for their assistance: Angela, Aynsley, Claire, Lyn, Nicole, Nicki, Penny, and Thomas. To my collaborators from the Philippines, Cebu Technological University-Argao through Hemres Alburo and the Philippines Biodiversity Conservation Foundation Inc. through Lisa Paguntalan for being great partners in this project. To Val Salares of the University of San Carlos and to the staff of the Philippine University Herbarium at the University of the Philippines-Diliman for the help in preparing herbarium specimens. To the Department of Environment and Natural Resources in the Philippines and their regional counterparts in Bataan, Benguet, Cebu, La Union, and Leyte for the assistance in securing permits. To the local government, the community, and my field guides in Mariveles (Bataan), Itogon (Benguet), Alcoy (Cebu), Boljoon (Cebu) Dalaguete (Cebu), Cebu City (Cebu), Argao (Cebu), Bauang (La Union), and Dagami (Leyte) for accommodating me during fieldwork. To the Conservation, Systematics, and Evolutionary Research Team aka ConSERT: Aisling, Ilina, Jana, Jasmine, Jennifer, Levi, Michael, Natalie, Rachel, Seoljong, and Stephanie, for fostering an environment of belongingness in the lab. To Craig and Heather, for the encouraging words. To Nina and the International Relations team for the help on scholarship matters. To my Filipino and Christchurch friends here and overseas for the support. To my mother and family back in the Philippines for the love. To Marcus's family, especially his parents, Steve and Chris, for making me feel at home in their house. To Marcus, for being always there especially during the last few weeks of my PhD study when it is the most challenging. Thanks for lending a listening ear to my endless rant about my PhD and for helping me take care of my mental health. I love you and I am lucky to have you.

ABSTRACT

Conservation genetic studies inform conservation management that aims to maintain genetic diversity for the long-term persistence of species. To my knowledge, published conservation genetic studies are currently lacking in Cebu, a Philippine island that has a long history of deforestation and has lost nearly all of its forest cover. Consequently, the effects of habitat fragmentation on patterns of genetic diversity and genetic connectivity among the remaining forests of Cebu remain unknown. As a first step towards filling this knowledge gap, microsatellite data from *Tetrastigma loheri* Gagnep. (Vitaceae), a commonly encountered woody vine species in the forests of Cebu, was used in Chapter 4 of this thesis to study patterns of genetic diversity and genetic connectivity among the four largest of the few remaining forested areas in Cebu. However, indications that Philippine *T. loheri* is a member of a species complex, referred to as the *T. loheri* s.l. complex, currently complicate this effort. In Chapters 2 and 3 of this thesis, I therefore used morphological and phylogenetic approaches to test the hypothesis that *T. loheri* s.l. is composed of more than one species. In Chapter 2, the results of unsupervised clustering analyses of geometric morphometric and other morphological datasets of vegetative characters revealed the absence of morphologically distinct clusters of individuals. In Chapter 3, several putative species were identified by species delimitation models from DNA sequence phylogenies of *T. loheri* s.l., but these were statistically poorly-supported and a supervised clustering method did not result in the identification of vegetative characters that characterise them. The results of the morphometric and phylogenetic analyses presented in Chapters 2 and 3 therefore do not provide conclusive evidence in support of recognizing more than one species within *T. loheri* s.l. in the Philippines. Even if this conclusion proves incorrect following future research, the results still suggest that the *T. loheri* specimens that were sampled from Cebu for the conservation genetic study in Chapter 4 are conspecific because these specimens were resolved as part of the same clades identified by species delimitation models as putative species. This means that the microsatellite dataset that I compiled for *T. loheri* samples from Cebu can be used for studying patterns of genetic diversity and genetic connectivity among the remaining forested areas in Cebu. The results of these analyses did not reveal evidence of low genetic diversity, despite suggesting a relatively high level of inbreeding in each of the four forested areas. Furthermore, low levels of genetic connectivity were evident among these areas, as inferred from identifying significant genetic differentiation among them. The negative consequences of habitat fragmentation, such as

inbreeding and low genetic connectivity, are likely to be greater for plant species that are less common than *T. loheri* and that have smaller population sizes or more restricted distributions. I therefore recommend the establishment of ecological corridors to increase genetic connectivity between the remaining forested areas with the aim of reducing the risk of inbreeding and loss of genetic diversity.

CHAPTER 1: General Introduction

1.1 THESIS OVERVIEW

This PhD thesis presents the results of my taxonomic and conservation genetics studies of *Tetrastigma loheri* Gagnep. (Vitaceae) in the Philippines. Chapter 1 introduces the background of my thesis project, outlines its general research aims, and includes a broad discussion of the methodology and data used. The research in Chapters 2 and 3 aims to contribute to resolving the *T. loheri* s.l. species complex in the Philippines. In Chapter 2, I used geometric morphometrics and other morphological approaches to study patterns of leaf shape diversity and variation in other vegetative characters with the aim of identifying morphologically distinct species within the *T. loheri* s.l. complex. Chapter 3 presents a molecular phylogenetic approach in which I used model-based species delimitation methods in combination with supervised morphology analyses to resolve the complex. The research presented in Chapter 4 is a conservation genetic study that used microsatellite data of *T. loheri* to reveal patterns of genetic diversity and connectivity among four remaining forested areas in Cebu, a large island that has lost much of its natural forest habitat. Chapter 5 provides a summary of the major findings of this thesis and presents some areas for further research.

1.2 DEFORESTATION IN THE PHILIPPINES AND CEBU

The Philippines has lost much of its original forest cover (Bankoff, 2007; Laurance, 2007). When Spanish colonisers first arrived in the Philippines in the 16th century, an estimated 90% of the land area was covered with tropical rain forest (Westoby, 1989; Lasco et al., 2001). This declined to 70% by the end of the Spanish colonisation (late 19th century) and continued to decrease to 50% at the time of Philippine independence from the United States (Bankoff, 2007). Deforestation accelerated in the 1960s as a result of commercial logging (Food and Agriculture Organization, 1997). By the 1980s, forest cover had diminished to 24.7% (Acosta et al., 2006). Using data from 2015, the Forest Management Bureau (2019) estimated forest cover in the Philippines at 23%. The latter study defined forest as an area of at least 0.5 ha of trees that are at least 5 m in height and that has a canopy of which at least 10% is formed by trees. Using global forest data from Hansen et al. (2013), Global Forest Watch (2019) estimated that only 1.9% of the Philippine forest remained intact in 2016, defining intact forest as an unbroken forest ecosystem with no signs of human

activity and of a large enough size to maintain native biodiversity (Potapov et al., 2017). Forest in the Philippines is presently dominated by secondary vegetation that is mostly derived from residual forest at former logging sites and is experiencing different levels of degradation (Lasco et al., 2001).

Cebu, a large island situated in the central part of the Philippines (Visayas region; Figure 1.1), possibly faced the most extensive deforestation (Jakosalem et al., 2013). This island experienced early urbanisation during Spanish colonisation and was considered one of the most progressive islands in the Philippines during the early 19th century (Paguntalan et al., 2015). The early development of Cebu has been linked to the rapid decline of the forest in the island over the past centuries (Paguntalan et al., 2015). By 1875, only 6.6–11% forest cover remained in Cebu (Bankoff, 2007). Around the same period, Rabor (1959) reported that only small and isolated forest fragments could be found. According to the Forest Management Bureau (2019), Cebu's forest cover was 1.57% in 2015. Forest habitat is now mostly confined to small areas that are inaccessible to agriculture and wood harvesting (Paguntalan et al., 2015). Only eight significantly-sized forested areas remain, of which the largest is 1,036 ha (Nug-as forest in Alcoy and Boljoon) and only one enjoys formal protective status (the Central Cebu Protected Landscape; Figure 1.1).

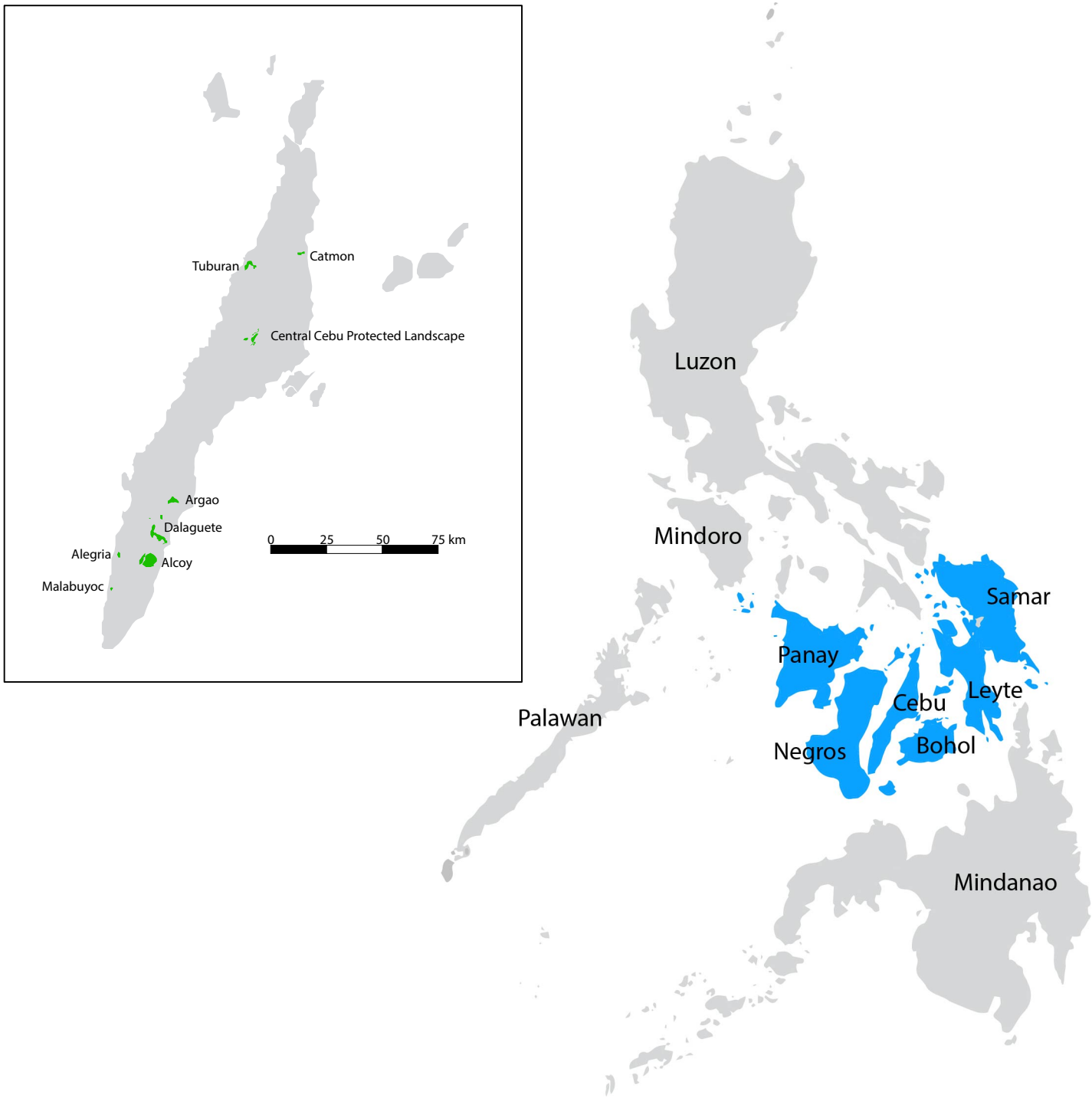


Figure 1.1 Map of the Philippines showing the location of select large islands. Islands in blue are part of the Visayas region. Lefthand inset shows a map with the remaining forested areas on the island of Cebu (redrawn from Paguntalan et al. (2015) with permission from the authors).

1.3 BIODIVERSITY AND CONSERVATION RESEARCH IN CEBU

The extensive deforestation on the island of Cebu has resulted in the extinction of many of its endemic wildlife, such as the Visayan warty pig (*Sus cebrifrons*) and Visayan spotted deer (*Rusa alfredi*) and many birds, such as the Cebu bar-bellied cuckoo shrike (*Coracina striata cebuensis*), Cebu Philippine oriole (*Oriolus steeri assimilis*), and Cebu white-bellied woodpecker (*Dryocopus javensis cebuensis*) (Paguntalan et al., 2015). Nonetheless, despite its fragmentation and degradation, Cebu's forest habitat remains to be home to its surviving wildlife, which includes fish, butterflies, damselflies, skinks, plants, birds, and bats (Gonzalez et al., 1999). Many of these are only found in Cebu. The island is particularly well known for its bird diversity. It is a centre of bird endemism in the Philippines, supporting a high number (at least 13) endemic bird taxa (Gonzalez et al., 1999; Paguntalan et al., 2015). Some of its remaining forested areas are designated as Important Bird Areas and Endemic Bird Areas (Paguntalan et al., 2015). In recent decades, a renewed interest in the biodiversity of Cebu has led to discovery of presumed extinct bats and birds (e.g., Dutson et al., 1993; Magsalay, 1993; Gadiana, 2004). Recently, new endemic plant species were described from the remaining forests of Cebu, such as the *Cynometra cebuensis* F.Seid. (Seidenschwarz, 2013), *Lepeostegeres cebuensis* Barcelona, Nickrent & Pelser (Pelser et al., 2016c) and *Vaccinium cebuense* Salares & Pelser (Salares et al., 2018). The presence of endemic and threatened biodiversity in the remaining forests in Cebu indicates the urgent need to conserve these forests and the biodiversity they support.

In recent years, there has been a growing number of conservation-related studies in Cebu. Some recent examples of conservation research include ecological studies of and breeding programs for threatened birds and bats in Cebu (e.g., Paguntalan & Jakosalem, 2008b; Jakosalem et al., 2013; Malaki et al., 2013; Paguntalan et al., 2015; Malaki, 2016), and biodiversity surveys of birds and plants (e.g., Lillo & Buot Jr; Paguntalan & Jakosalem, 2008a; Cadiz & Buot Jr, 2009; Lillo et al., 2015). However, while these studies are extremely valuable, there are to my knowledge no published studies of the genetic diversity of forest species in Cebu. Consequently, nothing is known about patterns of genetic diversity and connectivity for the few remaining forested areas in Cebu (Figure 1.1).

1.4 GENETIC DIVERSITY VS. SPECIES DIVERSITY

One of the reasons why genetic diversity has thus far not been explicitly incorporated in conservation research in Cebu might be the assumption that species diversity can serve as a surrogate for genetic diversity (Taberlet et al., 2012). Genetic diversity is the overall genetic variation within and among individuals of a species in a particular locality (Frankham et al., 2010), while species diversity refers to the number and abundance of species (Hamilton, 2005). Since both species diversity and genetic diversity are influenced by the same environmental factors, such as habitat size, connectivity, and variation in the environment, species and genetic diversity are often assumed to be correlated (Kahilainen et al., 2014). Although some studies indeed provide support for this correlation (e.g., Struebig et al., 2011; Blum et al., 2012), other research demonstrates that it does not always exist (Taberlet et al., 2012; Kahilainen et al., 2014). Kahilainen et al. (2014) argued that species and genetic diversity most likely only display parallel behaviour if species richness is influenced by the size of the habitat or by connectivity between populations. However, if species diversity is influenced by variation in the environment, then it is not a suitable proxy for genetic diversity, because variation in the environment affects species and genetic diversity in different ways. For instance, a heterogeneous environment can provide many rare habitats. Such an environment is expected to support high species diversity (Kahilainen et al., 2014). However, an increase in the number of rare habitats can also imply a reduction in the size of the habitat of each individual species, thus resulting in smaller populations. Small populations are in danger of genetic diversity loss through genetic drift. Therefore, an environment that is high in species richness, might be low in genetic diversity.

1.5 THE IMPORTANCE OF GENETIC DIVERSITY AND CONNECTIVITY IN CONSERVATION STUDIES

Genetic diversity is important for the long-term persistence of species, because it provides populations with a source of alleles that enables them to adapt to changes in the environment, thus increasing their evolutionary potential (Frankham et al., 2010). Populations that have high genetic diversity, therefore, have higher chances of long-term survival (Jamieson et al., 2008). Genetic diversity is influenced by genetic connectivity. Genetic connectivity reflects the degree to which gametes (i.e. pollen in plants) or seeds disperse between populations (Slatkin, 1987; Lowe & Allendorf, 2010). Genetic connectivity maintains genetic diversity by preventing loss of genetic diversity through genetic drift and

inbreeding (Ellstrand & Elam, 1993; Jenkins & Stevens, 2018). Because of their importance to the long-term persistence of species, genetic diversity and connectivity research, therefore, should, when possible, be explicitly incorporated in biodiversity conservation studies and be used to inform conservation management. Otherwise, because species diversity is not always a suitable proxy for genetic diversity, species-diversity based conservation research might either underestimate or overestimate genetic diversity, potentially resulting in conservation recommendations that might not be able to effectively increase the chance of long-term persistence of species.

1.6 PATTERNS OF GENETIC DIVERSITY AND GENETIC CONNECTIVITY IN FRAGMENTED HABITATS

Fragmentation of habitats causes contraction of continuous habitat into smaller isolated remnants, resulting in lower population sizes and increased isolation of populations (Young et al., 1996; Lienert, 2004). Small isolated populations are prone to genetic drift, because a population of few individuals has a high chance of losing alleles, especially rare ones, by chance (Frankham et al., 2010). This is especially likely to occur in small remnant populations, because only a subsample of parental alleles is transferred to the next generation (Ellstrand & Elam, 1993; Frankham et al., 2010). Loss of alleles can lead to fewer heterozygotes (individuals that contain two or more different alleles at a locus) and the eventual fixation of remnant alleles within the population (Ellstrand & Elam, 1993). Thus, genetic drift leads to loss of genetic diversity and consequently decreases the chance of long-term persistence of populations, because of the scarcity of alleles that would potentially allow them to adapt to a changing environment. In addition, in a small isolated population, offspring are likely to be produced as a result of mating between related individuals. This is known as inbreeding and may lead to inbreeding depression. Inbreeding increases the proportion of individuals in a population that carry two copies of the same allele at a locus (homozygotes) and this may expose deleterious recessive alleles which can reduce the reproductive fitness of individuals (i.e., inbreeding depression; Ellstrand and Elam, 1993) and therefore negatively affect the viability of populations (Young et al., 1996).

1.7 THE NEED FOR GENETIC DIVERSITY AND CONNECTIVITY DATA FOR CEBU'S FORESTS

The long history of deforestation in Cebu has resulted in the fragmentation of its forest habitat. Mostly likely, this has resulted in the decline and isolation of populations of forest species. Although this suggests that Cebu's forests are experiencing the negative consequences of forest fragmentation, such as inbreeding, genetic diversity loss, and low genetic connectivity among populations, there is to my knowledge, no published information about patterns of genetic diversity and genetic connectivity for its remaining forested areas. These data are important for informing conservation management that aims to maintain genetic diversity for the long-term persistence of species. For example, they enable conservation managers to prioritize areas that have the lowest levels of genetic diversity and connectivity. In acknowledgement of this knowledge gap, I investigated the current patterns of genetic diversity and connectivity for the four largest remaining forested areas in Cebu using genetic data from one of its forest species, *Tetrastigma loheri* Gagnep. (Chapter 4). *Tetrastigma loheri* is a woody vine from the grape family (Vitaceae) and it is a commonly encountered species in forest habitats in Cebu. It is dependent on forest vegetation, on which it climbs to reach the canopy and obtain sunlight. The reduction of forest cover resulting in smaller forest areas in Cebu is expected to have had an impact on the population sizes of *T. loheri* and the genetic connectivity of its populations. *Tetrastigma loheri* is abundant in all four selected forest areas and can therefore provide data about genetic diversity and connectivity among all four areas, as well as a sufficient number of individuals per area to yield statistically well-supported genetic patterns. This information can be used to develop area-specific hypotheses of patterns of genetic diversity and connectivity for other forest plant species.

1.8 DOCUMENTING GENETIC DIVERSITY USING MICROSATELLITE MARKERS

Microsatellite data was used to describe the genetic diversity within and among populations of *T. loheri* and the genetic connectivity among them. Microsatellites consist of short tandem repeats of 2–6 nucleotides and are distributed abundantly throughout the nuclear genome of most plants (Avise, 2004). They are highly polymorphic because of their high mutation rate and, thus, are a rich source of genetic variation for population genetic studies (Duminil et al., 2012). Microsatellites show co-dominance. Heterozygotes can

therefore be detected and allele frequencies can easily be estimated. They are genetic markers that are usually neutral. Neutral markers, such as microsatellites, are not influenced by selection and are independent of environmental conditions (Holderegger et al., 2006; Hall & Beissinger, 2014). In contrast, adaptive markers correspond to genetic variation that is under natural selection (Holderegger et al., 2006). Although neutral markers cannot be surrogates for adaptive markers, (Holderegger et al., 2006), neutral markers like microsatellites can be used to investigate genetic diversity and connectivity (Holderegger et al., 2006; Hall and Beissinger, 2014). Microsatellite studies are considered as a powerful and practical source of data in population genetics (Frankham et al., 2010). They have amongst others been successfully used to reveal genetic variation in forest plant species with the aim of informing their conservation management (e.g. *Carapa guianensis* Aubl., Dayanandan et al., 1999; *Caryocar brasiliense* Cambess., Collevatti et al., 2001; *Swietenia macrophylla* King, Lemes et al., 2003), including in the Philippines (e.g., *Rafflesia*; Pelser et al., 2017, 2018).

1.9 ESTIMATING GENETIC DIVERSITY

Genetic diversity was described in this thesis in terms of polymorphism of loci, allelic richness, heterozygosity, and the fixation index (a.k.a. inbreeding coefficient; F_{IS}). The polymorphism of loci refers to the presence of more than one allele at a locus. A locus is usually said to be polymorphic if it has the most frequent allele at a frequency of less than either 0.95 or 0.99 (Frankham et al., 2010). Allelic richness describes the number of alleles per locus in a population and should be only compared between populations when the sample sizes are similar (Frankham et al., 2010). Rarefaction can be applied to correct for the differences in sample size between populations so they can be comparable (Kalinowski, 2005). Expected heterozygosity (He) refers to the "heterozygosity expected for a randomly mating population with the given allele frequencies according to the Hardy-Weinberg equilibrium" (Frankham et al., 2010). Hardy-Weinberg equilibrium assumes that allele and genotype frequencies reach equilibrium and remain unchanged over generations in a large population where mating is random and there is an absence of mutation, migration and selection. It provides a null hypothesis against which the observed proportion of heterozygotes in a population (Ho) can be tested. Ho and He values that are not the same indicate violation of any of the assumptions of Hardy-Weinberg equilibrium. Thus, for instance, finding populations that have lower heterozygosity than expected by the Hardy-Weinberg equilibrium indicates decreased heterozygosity and increased homozygosity. Such

populations may be experiencing genetic drift or inbreeding. These particular consequences are expected for fragmented populations that have reduced population sizes with limited gene flow among them (Young et al., 1996). This could result in isolated populations that are genetically different from each other. The fixation index (F_{IS}) is a component of F-statistics (Weir and Cockerham, 1984) that is used to describe the degree of genetic differentiation within populations. F_{IS} is an index derived from comparing He and Ho and is used to determine the presence of inbreeding within the population. F_{IS} is expected to be greater than zero if inbreeding is present within a population and zero if inbreeding is absent.

1.10 ESTIMATING GENETIC CONNECTIVITY

Genetic connectivity can be measured in direct or indirect ways. Direct methods involve assigning individuals to their parents or subpopulation of origin based on their multiple-locus genotypes (Lowe & Allendorf, 2010). Indirect methods estimate genetic connectivity from the amount and pattern of genetic differentiation among populations (Lowe & Allendorf, 2010). Indirect methods using statistical approaches that estimate genetic differentiation among populations were used in this thesis: Analysis of Molecular Variance (AMOVA), a Bayesian clustering method using STRUCTURE, and Discriminant Analysis of Principal Components (DAPC).

Analysis of Molecular Variance (Excoffier et al., 1992) computes and tests the statistical significance of F_{ST}, a component of F-statistics that is used to measure the overall genetic differentiation among populations (global F_{ST}) or between populations (pairwise F_{ST}). Populations that are genetically differentiated because of a lack of genetic connectivity will have an F_{ST} estimates significantly greater than zero. Despite its utility in estimating genetic differentiation among populations, F_{ST} is sensitive to the total genetic variation among individuals and when genetic variation among individuals is high among genetically differentiated populations, it can lead to low but significant estimates of F_{ST} (Hedrick, 2005; Meirmans and Hedrick, 2011). The standardised F_{ST}, or simply F'_{ST}, is defined as the proportion of the maximum F_{ST} that can be obtained for the level of genetic variation present within a population (Hedrick, 2005; Meirmans and Hedrick, 2011). F'_{ST} allows comparison between loci with different levels of variation and between species with different effective population sizes.

STRUCTURE (Pritchard et al., 2000) is a commonly used program that implements a Bayesian clustering method for population genetic structure analysis (e.g., Ballian et al.,

2006; Bergl & Vigilant, 2007; Craft & Ashley, 2007; Aizawa et al., 2009; Ouinsavi et al., 2009; Dubreuil et al., 2010; González-Martínez et al., 2010; Debout et al., 2011; Wang et al., 2011; Reddy et al., 2012; Logan et al., 2015; Goncalves et al., 2019; Stojnić et al., 2019; Tamaki et al., 2019; Vaishnav et al., 2019). STRUCTURE is a model-based clustering program and identifies genetic clusters that minimize Hardy-Weinberg and linkage disequilibrium. STRUCTURE employs a Bayesian iterative algorithm that analyses the distribution of genetic variation among populations and places samples into genetic clusters whose members share a similar pattern of variation (Porras-Hurtado et al., 2013). Individuals are assigned to the cluster in which they obtain the highest posterior probability (Safner et al., 2011).

Discriminant Analysis of Principal Components (DAPC) is a non-model genetic clustering method which utilises a k-means algorithm for finding the number of clusters in the dataset. Unlike STRUCTURE, DAPC does not assume Hardy-Weinberg and linkage equilibrium when identifying genetic clusters (Jombart et al., 2010). Because of this, DAPC is usually applied in studies of populations that are not in Hardy-Weinberg equilibrium such as those of crops that are highly inbred and experience non-random mating (e.g., Matos et al., 2013; Filippi et al., 2015; Campoy et al., 2016). DAPC is similar to Principal Component Analysis (PCA) in that they are both multivariate methods that aim to summarize variation into a reduced number of dimensions (Jombart et al., 2010). However, unlike PCA, DAPC can assess the number of clusters using k-means clustering. Furthermore, DAPC has a better way of visualizing the variation between clusters by maximizing the between-group differences while minimizing the within-group differences (Jombart et al., 2010).

1.11 SPECIES COMPLEXES AND CONSERVATION

Species delimitation is one of the core challenges of taxonomy. It requires adopting a definition of what ‘a species’ is. There are over 22 such species concepts (Claridge et al., 1997). The most prevalent definition of species is the biological species concept of Mayr (1942), which defines a species as a group of naturally interbreeding populations that are reproductively isolated from other such groups (Ereshefsky, 2007). In plant taxonomy, defining a species using the biological species concept can be difficult and usually impractical because in most cases, information about interbreeding is lacking when new species are discovered and described. Therefore, new species are instead typically recognized by morphological characters that set them apart from other species (e.g., Pelser et al., 2016a).

Thus, plant species are often delineated using a morphological species concept (Cronquist, 1978). One problem with using morphological characters for species delimitation is that they are sometimes incapable of distinguishing closely related species, because characters states can overlap between them, particularly in a group of species that form a species complex (e.g., *Primulina* species in China; Yang et al., 2019). A species complex can be defined as a group of closely related individuals that form an unknown number of species, due to species boundaries that are presently unclear. Amongst others, this can be the result of recent phylogenetic divergence, introgression, high phenotypic plasticity, and partial barriers to gene flow between ecotypes (Duminil et al., 2012). Recent phylogenetic divergence can result in a related groups of individuals that are genetically differentiated but have little morphological differentiation among them (Harrison & Larson, 2014). Using morphology to identify species boundaries between these groups might therefore be difficult. Introgression is the integration of genetic material from one species into another through repeated backcrossing of hybrids with their parents (Baack & Rieseberg, 2007), potentially resulting in the formation of a gradient of intermediate phenotypes that can confuse morphological distinction between these species. Phenotypic plasticity allows plant species to change their morphology in response to environmental conditions and helps them to adapt to these (Nicotra et al., 2010). Consequently, some of populations may look morphologically different despite not being genetically distinct, resulting in their incorrect recognition as separate species. Ecotypes are genetically distinct races of a species that occur as a result of different adaptations to heterogeneous environments created by differences in elevation, soil moisture, light intensity, soil concentration of heavy metals, etc. (Hufford et al., 2003). Heterogeneity in the environment can create partial barriers to gene flow between populations and imposes selection pressures on them. This can result in genetic heterogeneity with associated morphological differentiation, such as in leaf morphology, seed characters, phenology, or physiological activity (Linhart & Grant, 1996). Ecotype populations can have subtle morphological differences between them, causing partially isolated populations to be mistaken as distinct species.

Species complexes can cause problems in conservation management (Frankham et al., 2010). They have confounded taxonomists in distinguishing closely related species, resulting in the incorrect assignment of taxonomic status (Frankham et al., 2010). One crucial conservation implication of misidentification of a species is the denial of protection to an unrecognized, endangered species (e.g., *Helianthus exilis*; Rieseberg, 1991). Furthermore, when misidentification in conservation genetics research results in a species being

overlooked, resulting in datasets that are mistakenly composed of data of more than one species, erroneous genetic patterns may emerge. Consequently, inaccurate conservation recommendations could be implemented. The potential problems that species complexes might cause in conservation research indicate the need for resolving the taxonomic delimitation of species complexes that are of conservation interest.

1.12 THE *T. LOHERI* S.L. SPECIES COMPLEX

Tetrastigma loheri was identified for the research presented in this thesis as a suitable species for obtaining a first indication of patterns of genetic diversity and connectivity among the few remaining forested areas in Cebu (Chapter 4), because it is present in all four areas selected for the study and can therefore provide data about the genetic connectivity among all these areas. In addition, this species is common enough to be able to provide a suitable number of samples from each area to provide statistically well-supported genetic patterns (Hale et al., 2012). However, some problems with identifying the species boundaries of *T. loheri* in the Philippines came to light in a previous study (Pelser et al., 2016b). This study revealed that plants identified as *T. loheri* in the Philippines display large morphological variation in leaf morphology, such as in the size, shape, and length/width ratio of the leaves and leaflets, the shape of the leaflet margin, venation patterns of leaflets, and in the length of petioles and petiolules (Pelser et al., 2016b). It is unclear if morphological variants are distinct species and if *T. loheri* is part of a species complex (i.e. the *T. loheri* sensu lato species complex; Pelser et al., 2016b), or if *T. loheri* merely displays intraspecific variation as a result of phenotypic plasticity or local adaptation. Other observations by Pelser et al. (2016b) also indicated the need for a taxonomic study aimed at clarifying the delimitation of *T. loheri*. For example, several individuals of *T. loheri* s.l. from the same island in the Philippines were found nested in different lineages in their DNA sequence phylogenies. This might suggest sympatry of reproductively isolated lineages representing different species, despite morphological similarities. Moreover, although considered taxonomically distinct when narrowly delimited, *T. loheri* (sensu stricto) shows morphological similarity with other previously and currently recognised *Tetrastigma* species in the Philippines: *T. diepenhorstii* (Miq.) Latiff, *T. philippinense* Merr., *T. stenophyllum* Merr., and *T. trifoliolatum* Merr. (Pelser et al., 2011 onwards). These species are different from *T. loheri* s.s. in relatively subtle morphological differences (as discussed in Chapter 2) and some *T. loheri* s.l. plants cannot be easily accommodated in any of these species based on the morphological

descriptions in their protologues (Miquel, 1861; Gagnepain, 1910; Merrill, 1912, 1914, 1916), because these *T. loheri* s.l. plants show a conflicting combination of character states.

The research presented in Chapters 2 and 3 aims to determine whether the *T. loheri* s.l. complex is comprised of more than one species, so that the delimitation of *T. loheri* is clarified for the conservation genetic study of Chapter 4.

To define species boundaries within *T. loheri* s.l., a unified species concept proposed by De Queiroz (2007) was used in Chapters 2 and 3. The unified species concept treats a species as a "separately evolving metapopulation lineage" that "acquires certain properties through the course of its divergence". De Queiroz proposed the unified species concept as an attempt to harmonise different competing species concepts (e.g. biological species concept of Mayr (1942), phylogenetic species concept of Baum and Shaw (1995), morphological species concept of Cronquist (1978)) in the literature without ignoring the important source of evidence that each species concept presents. For example, sympatric populations that form distinct lineages might not be considered as separate species by a morphological species concept (Cronquist, 1978) because the populations might be indistinguishable morphologically.

De Queiroz recognised a common element among the rival species concepts, i.e. that a species is a "separately evolving metapopulation lineage", which he argued is the only necessary property of a species. He further explained that secondary properties that are used to define a species such as monophyly, reproductive isolation, morphological diagnosability, coalescence of alleles, among many others, can be interpreted as contingent properties of a species acquired during the process of speciation. Under the framework of a unified species concept, the contingent properties, therefore, can serve as "important operation criteria or lines of evidence relevant to assessing the separation of lineages", and thus, are useful to "inferring the boundaries and numbers of species". Any contingent property, even a single one, is sufficient evidence to support the existence of a species. For instance, a species can be defined by morphology or it can be defined by both morphology and monophyly. Multiple lines of evidence present a highly-supported hypothesis of the existence of a species.

1.13 RESOLVING SPECIES COMPLEXES USING LEAF MORPHOLOGY

Although morphology might not always be able to conclusively differentiate species (e.g., cryptic species), reproductive and vegetative morphological features of plants are in many plant taxa helpful when distinguishing species (Cope et al., 2012). Leaf characters, for

example, are often used for taxonomic identification. Leaf morphology has been historically used to identify species in *Betula*, *Tilia, Ulmus,* and many more plant taxa (Cope et al., 2012) and to provide a better understanding of the taxonomy of genera such as *Gunnera* (Fuller and Hickey, 2005) and *Ticodendron* (Hickey and Taylor, 1991). Leaf morphology has previously also been used to inform phylogenetic relationships (Hickey and Wolfe, 1975), particularly before the advent of molecular phylogenetics. Among leaf characters, leaf shape is one of the most useful diagnostic features (Cope et al., 2012) and has the most discriminative power (Cope et al., 2012; Wäldchen & Mäder, 2018).

For many plant species, leaves are some of the most readily accessible parts of the plant for taxonomic study, because, in contrast to reproductive structures (e.g., flowers, inflorescences, fruits), they are available throughout the year or a substantial part of the year (e.g. deciduous species). Leaves are especially important for the taxonomic study of the *T. loheri* s.l. complex because flowers of *Tetrastigma* are rarely encountered in the field (Pelser et al., 2016b). Also in the fieldwork conducted for this study, only a few specimens of *T. loheri* s.l. with reproductive parts were found. This perhaps explains the scarce information in the literature about the flowers of *T. loheri* and related species (e.g., Merrill, 1914) and the lack of fertile herbarium specimens (Pelser et al., 2016b). In lieu of sufficient specimens with reproductive structures, vegetative characters provide the most accessible source of information for morphological taxonomic studies of the *T. loheri* s.l. complex. In Chapter 2, I aimed to resolve the *T. loheri* s.l. species complex in the Philippines by using geometric morphometric methods (discussed below) for finding groups that are different from each other in leaf shape. These could represent different species if a unified species concept is adopted using morphological diagnosability as evidence of lineage separation (De Queiroz, 2007).

1.14 GEOMETRIC MORPHOMETRICS

Morphometrics is the study of variation in shapes (Adams et al., 2004). Traditional approaches to morphometrics involve multivariate statistical analyses of measurements of length, width, and height of morphological structures (Adams et al., 2004). However, traditional morphometrics has been criticized for its limited power in the study of shapes. It lacks the capability to fully capture the geometry of the original structure, hence, some aspects of shapes are lost during the analysis (Rohlf & Marcus, 1993; Adams et al., 2004). Identifying homologous structures among samples can also be difficult, particularly, when

linear measurements are not based on homologous points, such as the maximum width of a structure (Adams et al., 2004). Furthermore, traditional morphometrics does not have standard correction methods for removing the effects of size and enabling the comparison of shapes among samples (Rohlf & Marcus, 1993; Adams et al., 2004).

In contrast to traditional morphometrics, geometric morphometrics can detect subtle differences between leaf shapes among samples. This approach captures the geometry of a morphological structure by placing anatomical landmarks or approximating the outline of a structure (Rohlf & Marcus, 1993; Adams et al., 2004; Viscosi & Cardini, 2011). Analysis of homologous points is achieved by comparing the relative positions of landmarks or the outlines of structures (Jensen, 2003; Viscosi & Cardini, 2011). The effect of non-shape variables such as size, position, orientation, or location are removed using methods like Procrustes alignment and Elliptic Fourier analysis (Adams et al., 2004; Cope et al., 2012). Landmark and outline methods are two popular geometric morphometric approaches.

The landmark method starts with the identification of homologous landmarks, which are biologically definable anatomical points in a structure (Jensen, 2003; Adams et al., 2004). Superimposition of these landmarks is then achieved through Procrustes alignment, ensuring that only the shape of the structure is retained and included in the analysis (Adams et al., 2004). One of the possible challenges when using the landmark method is finding a sufficient number of landmarks that can be used to capture the shape of a structure when only a few homologous points across samples are available (Jensen, 2003; Adams et al., 2004).

Outline analysis involves extracting information from the points of the outline of a structure and fitting these points with a mathematical function like in an Elliptic Fourier analysis (Adams et al., 2004). Elliptic Fourier analysis allows normalization to represent shapes independent of their orientation, size, or location (Cope et al., 2012). Shapes are then analysed by comparing the coefficients of the function as shape variables in a multivariate analysis (Adams et al., 2004). One criticism of outline analysis is that the set of points used to describe the shape outline does not necessarily have a one-to-one correspondence across samples (as in landmark method), although the outlines of the structures can be assumed as homologous (Adams et al., 2004).

Geometric morphometric methods have proved useful in taxonomic studies that used leaf shape data, including those that aimed to resolve taxonomic delimitations within morphologically diverse species complexes and in other taxonomically complicated groups (e.g., Nery & Fiaschi, 2019). For example, they have been used to distinguish individual species and hybrids of maples (Jensen et al., 2002), European oaks (Viscosi et al., 2009), as

well as two species and intraspecific genotypes of grapevines (Klein et al., 2017). Moreover, geometric methods have also been used for other plant parts than leaves (e.g., pollen: Bonhomme et al., 2013; lip petals of orchids: Shipunov & Bateman, 2005), demonstrating their wider use for studying shape variation.

1.15 CLUSTERING METHODS

Clustering (classification) methods can be used to find patterns of morphological similarity among and within putative species by analyzing geometric morphometric and other morphological datasets. The clustering methods used in this study of the *T. loheri* s.l. complex in the Philippines are grouped into two categories: unsupervised and supervised clustering methods. Chapter 2 employs a set of unsupervised clustering methods to analyse leaf shape data and a dataset composed of other vegetative morphological characters, whereas a supervised clustering method is used in Chapter 3 to determine if molecular phylogenetic clades are diagnosably distinct.

Unsupervised clustering methods do not use pre-defined labels that would indicate grouping properties in the dataset, hence, are suitable for identifying unknown groups (Frades & Matthiesen, 2010). Several unsupervised clustering methods were used in the research presented in Chapter 2: k-means cluster analysis, model-based clustering, and two ordination approaches: Principal Component Analysis (PCA) and Principal Coordinate Analysis (PCoA). K-means clustering aims to find clusters of specimens that meet the condition of having a minimum amount of variation within them and a maximum amount of variation among them. It used a k-means algorithm for this (MacQueen, 1967). Model-based clustering uses statistical models to describe the clusters and finds the best way to optimise the fit between these models and the observed data (Gan et al., 2007). PCA and PCoA are multivariate analysis techniques that visually summarize patterns of similarities among taxonomic units into a two- or three-dimensional ordination plot (Abdi & Williams, 2010). Unsupervised clustering methods are commonly applied in taxonomic studies (e.g., Petrini and Fisher (1988); Nikolić (1995); Joly and Bruneau (2007), Elisens and Nelson (1993); Joly and Bruneau (2007); Tjaden and Cohen (2006); Devos et al. (2007)).

On the other hand, supervised clustering methods classify samples into the pre-defined categories (Frades & Matthiesen, 2010). A Random Forest analysis, an example of a supervised clustering method, generates multiple decision (classification) trees from bootstrap samples of the original dataset and selects the tree that best classifies the samples

using the categories set in the analysis. A cross-validation test is performed by classifying the out-of-the bag (OOB) data (data that were not sampled during bootstrapping) using the decision trees generated by the analysis (Breiman, 2001; Liaw & Wiener, 2002). Also Random Forest analyses are commonly used in taxonomic studies (e.g., Shipunov et al., 2011; Skoracka et al., 2014; Moffat et al., 2015).

1.16 RESOLVING SPECIES COMPLEXES USING A PHYLOGENETIC APPROACH

The difficulty of distinguishing morphologically similar species can partially be overcome by using molecular genetic methods. Because molecular methods make use of markers that are by nature genetic, they are capable of detecting differentiation among species that might not be evident from morphology alone (Avise, 2004). Thus, molecular methods using data such as DNA sequences have become popular in species delimitation of species complexes (Vogler & Monaghan, 2007). DNA sequence data can be a source of highly informative genetic data in which every nucleotide position is treated as a character with four character states (i.e. A, T, C, G). These character states are unambiguous, unlike some morphological character states, which can be difficult to distinguish when they overlap. Furthermore, DNA sequences can provide myriads of genetic data points (Avise, 2004). For example, the commonly sequenced Internal Transcribed Spacer of the nuclear cistron (ITS) has around 800 nucleotides which potentially can provide several hundreds of informative characters. DNA sequence datasets therefore typically contain more characters than a corresponding morphometric dataset.

Because of their variability, plastid and nuclear DNA regions have been the most popular source of sequence data in phylogenetic studies of plant species. The use of sequence data from mitochondria is relatively uncommon in plants, because they are often not variable enough among species (Duminil & di Michele, 2009). The development of universal primers for some of the most commonly sequenced DNA regions (e.g. ITS) makes these relatively easy to sequence. Regions found in the intergenic spacers of the plastid genome, such as *atpB-rbcl, psbA-trnH, trnL-F*, and introns such as *rps16* and *trnL* (Chen et al., 2011; Wen et al., 2013; Pelser et al., 2016b) as well as the ITS of the nuclear genome (Pelser et al., 2016b) have been useful sources of data for determining species-level relationships in Philippine Vitaceae, including Philippine *Tetrastigma*. These DNA regions were therefore used in the species delimitation research outlined in Chapter 3.

Although thus far not used in studies of Vitaceae, the External Transcribed Spacer (ETS) is one of the most utilized nuclear DNA regions for phylogenetic study along with the ITS (Poczai and Hyvönen, 2010). Like the ITS, it is part of the ribosomal DNA, which is present in many copies in the nuclear genome. ETS is part of the intergenic spacer region between the tandem repeats of the ribosomal gene (18S-5.8S-26S). It consists of two sites: the 3'-ETS which borders the downstream part of the 26S and the 5'-ETS which borders the upstream regions of the 18S. Previous studies have demonstrated the utility of ETS for studying closely related species (e.g., Linder et al., 2000; Sallares & Brown, 2004). The ETS region has been shown to evolve faster than the ITS (Markos & Baldwin, 2001; Poczai & Hyvönen, 2010). Thus, ETS can potentially be more variable and more phylogenetically informative than ITS. When combined with ITS data, the ETS has been shown to improve resolution of ITS-based phylogenies (e.g., Baldwin & Markos, 1998; Bena et al., 1998; Markos & Baldwin, 2001; Saar et al., 2003; Calonje et al., 2009). In addition to ITS and *atpB-rbcL, psbA-trnH, rps16* , *trnL,* and *trnL-F* sequences, 5'-ETS data was therefore used for phylogeny reconstruction in Chapter 3.

Phylogenetic analysis of DNA sequence data can be used to provide insights about the descent of lineages through the reconstruction of their evolutionary history. A DNA sequence alignment is constructed, which forms the input of phylogenetic reconstruction analyses such as Maximum Parsimony (Fitch, 1971), Maximum Likelihood (Felsenstein, 1981) and Bayesian Inference (Huelsenbeck et al., 2001). These are used to generate hypotheses of evolutionary relationships in the form of phylogenetic trees. Amongst others, these trees can be used for identifying species boundaries. For example, Li & Yan (2013) used the results of a phylogenetic analysis of ITS and *matK* DNA sequences to provide evidence in support of the recognition of *Liparis pingxiangensis* L.Li & H.F.Yan as a new species. Recently, more formal approaches for phylogeny-based species delimitation have been developed. The Generalized Mixed Yule Coalescent (GMYC) and the Poisson Tree Processes (PTP) methods are two approaches that were used in the research presented in Chapter 3. They have been proven useful in exploring the presence of undetected species diversity within poorly delimited groups (e.g., Carstens et al., 2013). The GMYC and the PTP methods both estimate the interspecific processes (speciation) and intraspecific processes (coalescence into a population) along the branches of a phylogenetic tree using information from branching rates. GMYC estimates the time of the switch between speciation and coalescence on the branches (Fontaneto et al., 2007; Fujisawa & Barraclough , 2013; Pons et al., 2006). Putative species are marked by the nodes after that switch. On the other hand, PTP models speciation and

coalescence events directly from the branching rates (Zhang et al., 2013) and classify the branches into species and population level processes (Zhang et al., 2013; Tang et al., 2014). GMYC is thought to identify evolutionary processes more accurately than PTP, because speciation and coalescence happen in the context of time and not necessarily in relation to how many substitutions occur in a locus (Tang et al., 2014). However, GMYC relies on the accuracy of an ultrametric tree input and the process of generating an ultrametric tree from a phylogeny is computationally intensive and error-prone (Zhang et al., 2013). PTP instead uses non-ultrametric phylogenetic trees as an input (Zhang et al., 2013).

Species boundaries within the *T. loheri* s.l complex were determined in Chapter 3 by using GMYC and PTP species delimitation models on DNA sequence phylogenies. Furthermore, a supervised clustering method was employed to determine if the putative species resolved by these species delimitation models are diagnosably distinct. Such groups could be recognized as species under the unified species concept (De Queiroz, 2007) using monophyly and morphological diagnosability as evidence of lineage separation.

1.17 AIMS OF THIS THESIS

Nothing is presently known about the effects of habitat fragmentation on the genetic diversity of plant species in the remaining forests of Cebu, a Philippine island that has lost nearly all of its forest cover. To contribute to filling this knowledge gap, I studied patterns of genetic diversity and genetic connectivity for the four largest remaining forest areas in Cebu (Alcoy, Argao, Dalaguete, and the Central Cebu Protected Landscape (CCPL); Figure 1.1), using microsatellite data of *Tetrastigma loheri*, a commonly encountered forest vine. However, previous research indicates that this species might be part of a species complex: *T. loheri* s.l. Because this may complicate my conservation genetics study, I studied the species delimitation of *T. loheri* s.l. using unsupervised and supervised clustering analyses of a dataset of vegetative morphological characters, as well as species delimitation models on a DNA sequence phylogeny of the complex.

I specifically aimed to determine:

1. if *T. loheri* s.l. in the Philippines is composed of more than one species as indicated by morphological discontinuities among groups of specimens in leaf shape and other vegetative characters. (Chapter 2)

2. if *T. loheri* s.l. in the Philippines is composed of more than one species by identifying monophyletic groups in DNA sequence phylogenies that are identified as putative species by Generalized Mixed Yule Coalescent and Poisson Tree Processes species delimitation models and that are diagnosably distinct in leaf shape or other vegetative morphological characters. (Chapter 3)

3. patterns of genetic diversity and genetic connectivity among Alcoy, Argao, Dalaguete, and CCPL by answering the following questions (Chapter 4):

a. What is the genetic diversity of *T. loheri* in Alcoy, Argao, Dalaguete, and CCPL?
b. Is inbreeding evident in Alcoy, Argao, Dalaguete, and CCPL?
c. What is the pattern of genetic connectivity among the populations of *T. loheri* in Alcoy, Argao, Dalaguete, and CCPL?

CHAPTER 2: Morphological diversity in leaf shape and other vegetative characters suggests that the *Tetrastigma loheri* s.l. complex (Vitaceae) is not composed of more than one species.

2.1 ABSTRACT

The long history of deforestation in the Philippines has resulted in the fragmentation of its original forest habitats across the archipelago. The extent of the genetic consequences of forest fragmentation on plant populations is, however, largely unknown in Cebu, an island that is almost devoid of forest. As a first step towards filling this knowledge gap, microsatellite data from *Tetrastigma loheri* (Vitaceae), a commonly encountered woody vine species in the forests of Cebu, will be used in Chapter 4 of this thesis to study patterns of genetic diversity and genetic connectivity among four of the few remaining forested areas in Cebu. However, indications that *T. loheri* is comprised of more than one species in the Philippines currently complicate this effort. In the research presented in this Chapter, I therefore tested the hypothesis that *T. loheri*, as currently circumscribed for the Philippines, is composed of more than one species by performing a geometric morphometric study of leaf shape diversity and studying patterns of variation of a selection of other vegetative morphological characters. PCoA, PCA, k-means clustering and model-based clustering analyses of 97 Philippine specimens ascribed to *T. loheri*, including the morphologically similar *T. philippinense, T. trifoliolatum*, and *T. stenophyllum* (i.e. *T. loheri* sensu lato), suggest the absence of morphological discontinuities among these previously recognized taxa. These results therefore do not support the existence of more than one species within *T. loheri* s.l. in the Philippines. If correct, this means that *T. philippinense,* Philippine specimens of *T. trifoliolatum*, and *T. stenophyllum* are not taxonomically distinct from *T. loheri*.

2.2 INTRODUCTION

In the past few centuries, the Philippines has lost much of its forest cover (Bankoff, 2007; Laurance, 2007). Presently, according to a recent estimate of the Forest Management Bureau (2019), only 23% of forest cover remains in the Philippines. Forest destruction, degradation and fragmentation as a result of deforestation reduces population sizes of species that are dependent on natural forest for their existence (Young et al., 1996). It also increases the distance between their populations (Young et al., 1996) and this can lead to a loss of genetic connectivity among populations (Ricketts, 2001; Harris & Reed, 2002), resulting in

reduced gene flow among them. Reduced gene flow between small populations increases the risk of genetic diversity loss through genetic drift (Ellstrand & Elam, 1993; Young et al., 1996). However, maintaining genetic diversity is essential to the long-term persistence of species because it provides them with evolutionary potential by enabling populations to adapt to changes in their environment (Jamieson et al., 2008; Frankham et al., 2010). To our knowledge, detailed information about patterns of genetic diversity and connectivity of plant populations in fragmented forest habitats in the Philippines (e.g., Pelser et al., 2017; Pelser et al., 2018) is scarce. This is certainly the case in Cebu, a large island that is almost devoid of forest cover (Seidenschwarz, 1988; Gonzalez et al., 1999; Bensel, 2008; Paguntalan et al., 2015).

Tetrastigma loheri Gagnep. is a species of woody vines in the grape family (Vitaceae) that is native to Borneo and the Philippines (Pelser et al., 2011 onwards). It is commonly encountered in the primary and secondary forests of the Philippines (Pelser et al., 2016b). Because it is also a common forest species in Cebu, *T. loheri* may be a suitable species for studying patterns of genetic diversity and connectivity among its fragmented forest ecosystems (Chapter 4). However, the use of *T. loheri* for such conservation genetics research is complicated by taxonomic issues. In the Philippines, *T. loheri* forms a species complex with several other currently or previously recognized species (i.e. the *T. loheri* sensu lato complex; Pelser et al., 2016b). A species complex is a group of closely related populations that form an unknown number of species, due to species boundaries that are presently unclear. Amongst others, it can be comprised of morphologically different populations that are genetically similar (e.g., *Trollium* species; Després et al., 2003) or can consist of geographically proximate and morphologically similar populations that form distinct genetic lineages (e.g., *Mercurialis annua*; Ma et al., 2019). There are several indications that *T. loheri* s.l. might be composed of more than one species in the Philippines. Firstly, *T. loheri* s.l. has a broad range of variation in its vegetative morphology. It is unclear if these morphological variants are distinct species or if they merely display intraspecific variation as a result of phenotypic plasticity or local adaptation. Secondly, several individuals of *T. loheri* s.l. from the same island in the Philippines were found nested in different lineages in a previous molecular phylogenetic analysis (Pelser et al., 2016b). This might indicate sympatry of reproductively isolated lineages representing different species despite morphological similarities.

When most narrowly delimited, *T. loheri* (sensu stricto) is considered taxonomically distinct from four species of the *T. loheri* s.l. complex that have been recognized for the

Philippines, i.e. *T. diepenhorstii* (Miq.) Latiff*, T. trifoliolatum* Merr.*, T. philippinense* Merr.*,* and *T. stenophyllum* Merr. (Pelser et al., 2011 onwards), but the morphology of their type specimens and their respective protologues indicate that all are similar in overall appearance, particularly by sharing glabrous and coriaceous leaves that almost always have three leaflets and never more than that, in combination with petiolules that are longer on the terminal leaflet than on the lateral leaflets, and a leaflet margin that is generally toothed (Miquel, 1861; Gagnepain, 1910; Merrill, 1912, 1914, 1916). *Tetrastigma diepenhorstii, T. trifoliolatum, T. philippinense,* and *T. stenophyllum* only appear to be different from *T. loheri* s.s. in relatively subtle morphological features.

Latiff (2001) concluded from a comparison of their type specimens that *T. diepenhorstii* (originally described from Sumatra; Miquel 1861) and *T. trifoliolatum* (from the Philippines; Merrill, 1914) are synonymous and listed their combined distribution to comprise Sumatra, Borneo, and the Philippines. This view was adopted by Zakaria et al. (2016, 2017). Chen et al. (2011) however concluded from a molecular phylogenetic study in which they included specimens that they identified as *T. diepenhorstii* and *T. trifoliolatum* that these names represent different species. Although the Philippine specimen that Chen et al. (2011) identified as *T. diepenhorstii* was resolved as nested in a clade of Philippine *T. loheri* s.l. specimens in a subsequent phylogenetic study (Pelser et al., 2016b), the exclusively non-Philippine samples of *T. trifoliolatum* sequenced by Chen et al. (2011) were more closely related to other *Tetrastigma* species than *T. loheri* s.l. in Pelser et al. (2016b).

Both *T. diepenhorstii* and *T. trifoliolatum* are recorded as having pubescent inflorescences, whereas those of *T. loheri* s.s. are reported to be glabrous (Miquel, 1861; Gagnepain, 1910; Merrill, 1914). In addition, *T. diepenhorstii* has elliptic to ovate (vs. lanceolate to oblong) leaflets and longer petioles (6.5–28 cm vs. up to 4 cm) than *T. loheri* s.s.(Miquel, 1861; Gagnepain, 1910; Zakaria et al., 2016, 2017). As described in its protologue (Merrill 1914), *T. trifoliolatum* has larger leaflets (14–20 x 7–10 cm vs. 4.5–12 x 2.5–5 cm), longer petioles (c. 12 cm vs. up to 4 cm), and longer petiolules of the lateral leaflets (3–4 cm vs. up to 0.5–1 cm), than *T. loheri* s.s. (Gagnepain, 1910).

Merrill (1912) described the Philippine endemic *T. philippinense* as having pubescent petals whereas those of *T. loheri* s.s. are glabrous according to Gagnepain (1910) and (Merrill, 1912). However, Merrill (1916) later synonymized *T. philippinense* under *T. loheri*. Although he did this without providing a detailed argument, this suggests that he considered their morphological differences too minor to merit taxonomic distinction at the level of species.

Tetrastigma stenophyllum is a taxon that is endemic to the Philippines. It reportedly has narrower leaves than *T. loheri* s.s. (Merrill, 1916). Merrill also mentioned that *T. stenophyllum* differs from *T. loheri* s.s. in having pubescent ovaries and stellate stigma lobes but did not explicitly mention the presence or type of indumentum of the ovaries of *T. loheri* nor the details of the morphology of its stigma lobes.

The original descriptions of the aforementioned members of the *T. loheri* s.l. complex were based on a very few specimens. Examination of a larger number of Philippine specimens resulting from a previous study (Pelser et al., 2016b) and my own fieldwork showed that there are plants that cannot be easily accommodated in any of these species, because they show a conflicting combination of character states. There is therefore a need for a study aimed at resolving the taxonomic delimitation of *T. loheri* s.l. This is particularly important in light of Chapter 4 of my thesis, because if *T. loheri* s.l. is composed of more than one species and if this is overlooked, it is possible that *T. loheri* datasets aimed at understanding patterns of genetic diversity and connectivity are unintentionally composed of data from different species. This would result in erroneous patterns and potentially incorrect conservation recommendations if *T. loheri* is used for investigating patterns of genetic diversity and connectivity between forest areas in Cebu.

The present chapter is the first part of a two-fold approach to resolving the *T. loheri* s.l. species complex. It aims to analyse leaf geometric morphometric and other vegetative morphological datasets using unsupervised clustering methods to test the hypothesis that the *T. loheri* s.l. group is composed of more than one group of individuals that are morphologically distinct in leaf shape and other vegetative characters. Finding such groups would support the presence of more than one species in *T. loheri* s.l. when a unified species concept is applied (De Queiroz, 2007) using morphological distinction as evidence of lineage separation. In Chapter 3, I will use the combined leaf geometric morphometric and other vegetative morphological datasets as part of a molecular phylogenetic approach to resolving the species complex. A supervised clustering method will then be applied to determine if this combined morphology dataset can discriminate putative species within *T. loheri* s.l. that are identified using phylogeny-based species delimitation models.

Vegetative characters provide the most accessible source of information for taxonomic studies in *T. loheri* s.l., because reproductive characters of *T. loheri* s.l. are poorly known. In part, this is because most herbarium specimens of *Tetrastigma* specimens that are available for study lack flowers or fruits (Pelser et al., 2016b). This is perhaps because flowering and fruiting plants are rarely encountered, or because these reproductive parts

might be difficult to see if these are primarily produced on parts of the vines that grow in the canopy of dense forest vegetation (Pelser et al., 2016b). Furthermore, *Tetrastigma* species are dioecious and staminate and pistillate flowers are therefore present on different individuals. This complicates species delimitation studies that use morphological approaches, because it can be difficult to determine if staminate and pistillate plants belong to the same species if differences in vegetative morphology are wanting. The lack of data about the diversity of reproductive morphological characters of *Tetrastigma* species presents a significant hurdle in taxonomic research on *Tetrastigma* and might be one of the reasons why a comprehensive taxonomic revision of this genus has not yet been produced. In this study, I therefore used vegetative characters as a source of potentially informative morphological characters for resolving the *T. loheri* s.l. complex.

Leaves of *T. loheri* s.l. plants vary in size, shape, venation, and length of their petiole and petiolules (Pelser et al., 2016b). Leaf characters might therefore be informative for identifying putative species in the species complex. Leaf shape has been considered to be one of the most useful diagnostic features in taxonomy (Cope et al., 2012). It has the most discriminative power among leaf characters because species often have characteristic leaf shapes (Cope et al., 2012; Wäldchen & Mäder, 2018). Furthermore, leaf shape is more determined, in many cases, by genetic information than by the environment (Cope et al., 2012), hence, it is a relatively stable character, unlike other characters such as leaf size, which can be different in different environmental conditions (e.g., Xu et al., 2009).

Geometric morphometrics can be used to detect subtle differences between leaf shapes among samples. It is capable of capturing leaf geometry by placing anatomical landmarks on leaves or by tracing their outline (Rohlf & Marcus, 1993; Adams et al., 2004; Viscosi & Cardini, 2011). It employs methods that remove the effect of size and other variables like leaf position, orientation, or location, ensuring that shape is exclusively accounted for in the analysis (Adams et al., 2004; Cope et al., 2012). Landmark and outline analyses are two commonly used geometric morphometric approaches (see examples of studies in Chapter 1). The landmark approach involves the identification and comparative analysis of the relative positions of biologically definable landmarks after their superimposition through Procrustes alignment (Adams et al., 2004). Procrustes alignment allows retention of only the shape information of a structure and generates shape coordinates that are used in a subsequent multivariate analysis (Adams et al., 2004; Viscosi & Cardini, 2011). Outline analysis involves digitising points along an outline of a structure and fitting these points with a mathematical function that normalises the data, such as an elliptic Fourier

analysis (Adams et al., 2004; Cope et al., 2012). The coefficients of the functions are used as shape variables and are compared in a multivariate analysis (Adams et al., 2004).

Because of their discriminatory power, both the landmark and outline methods will be used in this chapter to try and resolve the *T. loheri* s.l. complex by studying leaf shape variation. I will also study a selection of other vegetative morphological characters. I aim to test the hypothesis that *T. loheri* s.l. is composed of more than one species in the Philippines under the unified species concept (De Queiroz, 2007) assuming morphological distinction as a biological property acquired by a lineage in the process of speciation. Morphological discontinuities in leaf shape and other vegetative characters are here therefore considered as evidence for the presence of more than one species in the *T. loheri* s.l. complex.

2.3 OBJECTIVE

Using a geometric morphometric study of leaf shape and a morphological investigation of other vegetative characters, the present study aims to determine if *T. loheri* s.l. in the Philippines is composed of more than one species as indicated by morphological discontinuities among groups of specimens.

2.4 METHODOLOGY

2.4.1 Specimen sampling

A total of 97 herbarium specimens deposited in the University of Canterbury herbarium (CANU) were used for this study (Appendix 1). These specimens were collected at different localities in the Philippines to represent the range of distribution of *T. loheri* s.l. in this country. Seventy-nine specimens identified as *T. loheri* s.l. were collected during a previous study (Pelser et al., 2016b). The other 18 specimens were collected during fieldwork in November–December of 2016 and November 2017–January 2018 in areas that were not sampled by Pelser et al. (2016b). These additional specimens were deposited in CANU (University of Canterbury in New Zealand), CEBU (University of San Carlos in Cebu), and PUH (University of the Philippines in Quezon City) (Appendix 1).

In order to conclusively determine if previously described species are taxonomically distinct, their type specimens need to be included in a taxonomic study. Unfortunately, type specimens of *T. diepenhorstii, T. loheri*, *T. philippinense, T. trifoliolatum,* and *T. stenophyllum* were not available to me for this morphological study. In an effort to mitigate this and to increase the probability that, if existing, any morphologically distinct taxa within

the *T. loheri* s.l. complex in the Philippines would be discovered in this study, I collected specimens that most closely resembled *T. loheri*, *T. philippinense, T. trifoliolatum,* and *T. stenophyllum* from the type localities of these species. The type specimen of *T. diepenhorstii* is from Sumatra (Miquel 1861) and it was not possible for me to conduct fieldwork there.

2.4.2 **Geomorphometric analyses of leaf shape**

2.4.2.1 Specimen photography

The leaf shape morphology of *T. loheri* s.l. was studied using geomorphometric analyses of the terminal leaflets and the right-hand lateral leaflets (in abaxial view). Both the landmark and outline methods were used to record the morphology of the laminae of the leaflets, but only the landmark method also captured information about the petiole and petiolules. One representative leaf from each herbarium specimen was photographed for these analyses. Only relatively mature leaves with at least two leaflets in good condition, i.e. no or only minor damage, were selected. Leaves were photographed using a Nikon D3300 camera mounted on a camera stand with illumination from four light bulbs. All leaves were photographed to show the abaxial side, because this side shows the secondary veins most clearly. Some leaves were placed on a platform with an LED backlight during photography to make the secondary veins more obvious. An onion skin paper, which served as filter, was placed on the LED light to reduce brightness.

When necessary, leaflets that became detached from the remainder of the leaf during specimen preservation or handling were returned to their former position. For specimens with a missing or damaged right lateral leaflet, the left lateral leaflet was used. This was done by editing the photograph in Adobe Photoshop CC 19.1.6 by flipping the left leaflet over to assume the position of the right lateral leaflet (in other words, using the mirror image of the left lateral leaflet). For the few specimens with a missing leaflet apex, the shape of the leaf apex was redrawn in Photoshop by estimating the position of the leaflet apex and using it as a guide to complete the leaflet shape. Furthermore, a few of the specimens had some level of herbivory resulting in holes in the laminae. These holes usually extended to a portion of the margin of the leaflet. This was addressed by reconstructing the missing areas of the lamina by redrawing the leaflet margin and filling in the holes with color in Photoshop so that they would not be recorded as a distinguishing feature in the outline analyses.

2.4.2.2 Landmark method

2.4.2.2.1 Selection of landmarks

To capture the shape of the trifoliolate leaves of *T. loheri* s.l., a similar approach as that used by Klingenberg et al. (2012) and Chitwood & Otoni (2017) for selecting landmarks for compound leaves was used. Landmarks were placed at points on the leaf that would capture aspects of the shape of the compound leaf, including the shape, relative sizes, and position of the leaflets and the details of the leaflets (Figure 2.1). The descriptions of the 30 landmarks used in this study are provided in Table 2.1. These points are assumed homologous across specimens because they have a one-to-one correspondence in all specimens (Zelditch et al., 2004).

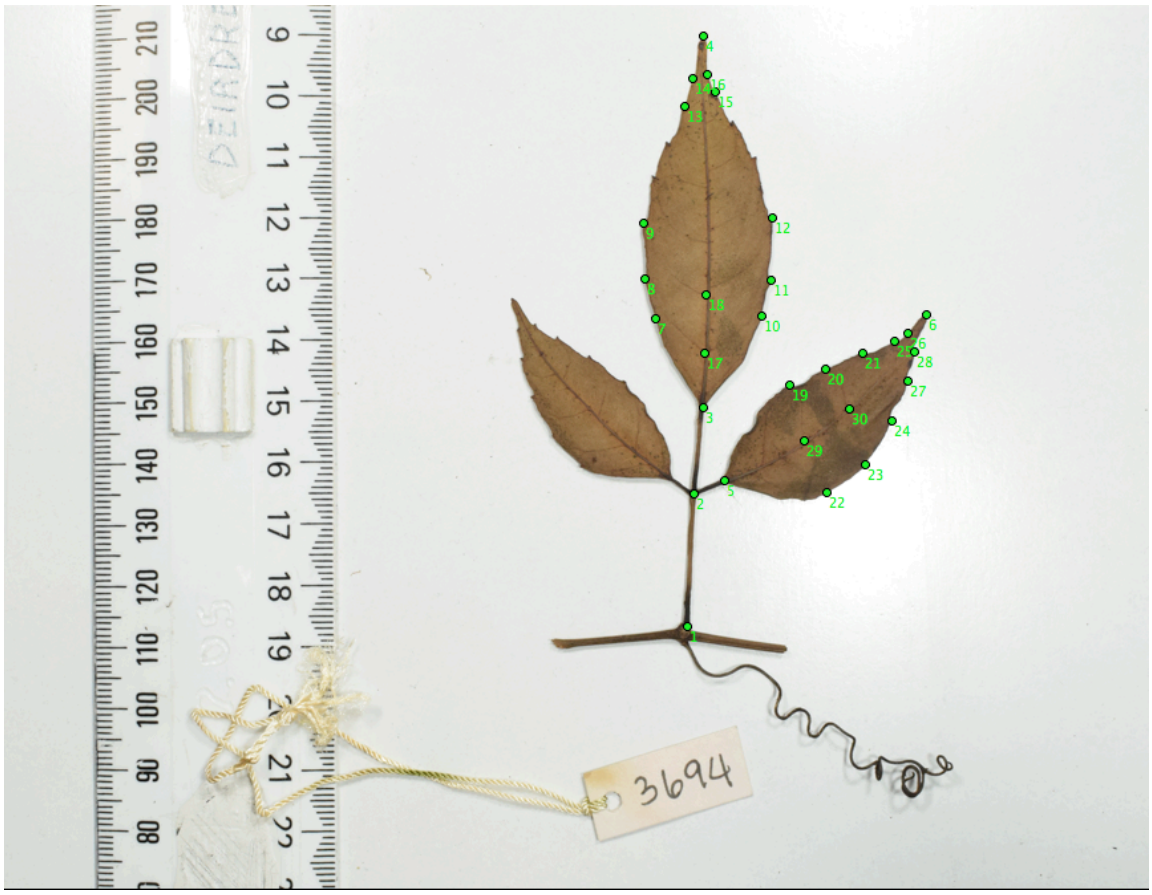


Figure 2.1. Location of landmarks on a leaf of a herbarium specimen of Tetrastigma loheri *s.l.*

Table 2.1. The number assignment and description of the landmarks used for a geomorphometric analysis of Tetrastigma loheri *s.l. leaf shape diversity.*

General leaf shape
1- The base of the petiole
2- The apex of the petiole
3- The apex of the petiolule of the terminal leaflet
4- The apex of the terminal leaflet
5- The apex of the petiolule of the lateral leaflet
6- The apex of the lateral leaflet
Terminal leaflet
Left half. Landmark points 7, 8, 9 were selected to record the shape of the base and greatest curvature of the lamina. Landmark points 4, 13, and 14 were assigned to record the morphology of the leaflet apex.
7- The first junction of the secondary veins and the teeth of the leaflet from the base
8- The second junction of the secondary veins and the teeth of the leaflet from the base
9- The third junction of the secondary veins and the teeth of the leaflet from the base
13- The second junction of the secondary veins and the teeth of the leaflet from the apex
14- The first junction of the secondary veins and the teeth of the leaflet from the apex
Right half. Not all terminal leaflets are symmetrical and to be able to record the shape of the entire lamina, similar landmarks to those on the left half of the leaflet were therefore placed on the right half.
10- The first junction of the secondary veins and the teeth of the leaflet from the base
11- The second junction of the secondary veins and the teeth of the leaflet from the base
12- The third junction of the secondary veins and the teeth of the leaflet from the base
15- The second junction of the secondary veins and the teeth of the leaflet from the apex
16- The first junction of the secondary veins and the teeth of the leaflet from the apex
Secondary veins and angles. The following landmarks were assigned to record variation in the secondary vein angles and the distance between two secondary veins from their points of origin at the midvein.
17- The point of origin of the second secondary vein from the base

18- The point of origin of the third secondary vein from the base
Lateral leaflet
Similar landmark points were placed on the lateral leaflets.
Left half.
19- The first junction of the secondary veins and the teeth of the leaflet from the base
20- The second junction of the secondary veins and the teeth of the leaflet from the base
21- The third junction of the secondary veins and the teeth of the leaflet from the base
25- The second junction of the secondary veins and the teeth of the leaflet from the apex
26- The first junction of the secondary veins and the teeth of the leaflet from the apex
Right half.
22- The first junction of the secondary veins and the teeth of the leaflet from the base
23- The second junction of the secondary veins and the teeth of the leaflet from the base
24- The third junction of the secondary veins and the teeth of the leaflet from the base
27- The second junction of the secondary veins and the teeth of the leaflet from the apex
28- The first junction of the secondary veins and the teeth of the leaflet from the apex
Secondary veins and angles.
29- The point of origin of the second secondary veins from the base
30- The point of origin of the third secondary veins from the base

2.4.2.2.2 Placement of landmarks and analysis

Photographs of leaves were imported in ImageJ version 1.52a (Abràmoff et al., 2004) to digitise the landmarks. The 30 landmarks were placed consistently in the same order in all specimens to enable comparison of homologous points. Coordinates were generated and were imported in MorphoJ version 1.06d (Klingenberg, 2011). Coordinates were then rescaled and centered at the origin using Procrustes analysis aligned using principal axes generating 60 Procrustes coordinates. A covariance matrix was then generated using these Procrustes coordinates.

2.4.2.2.3 Number of landmarks

To determine if the number of selected landmarks was sufficient for estimating the shape variation of the specimens used in this study, the function *lasec* (Watanabe, 2018) from the LaMBDA R package in R studio version 1.1.423 (RStudio Team, 2016) was used (www.github.com/akiopteryx/lambda). The resulting sampling curve plateaued and a diminishing variance in fit values in the plot was observed (Appendix 2). This was interpreted as evidence for stationarity in shape information and indicates that the 30 landmarks were adequate for characterising the shape variation of the specimens.

2.4.2.3 Outline method

The leaf photographs that were used for the landmark analysis were also used for the outline analysis. The terminal and lateral lamina of each specimen were cropped out separately and were converted to black images with white background using Photoshop. These photos were imported in Momocs version 1.2.9 (Bonhomme et al., 2014) in R studio using the function *import_jpg* which also extracted the outlines of the leaflet laminae from the photos. These outlines were converted to coordinates using the function *Out.* Two landmarks were placed on the leaf outlines to guide the alignment of the lamina shapes: one at the apex of the petiolule, and the other one at the apex of the lamina of the associated leaflet. This ensured that the shape outlines were all placed in the same orientation when superimposed (Appendices 3 & 4). A series of alignment and scaling functions suggested by Bonhomme (pers. com.) were used to further improve the alignment and standardise the shapes (Appendix 5). To characterise the shape of the outline, the aligned and scaled shapes were analysed with an elliptical Fourier analysis using the function *efourier*. The elliptic Fourier analysis uses harmonic functions to scan and obtain details of the outline (Cope et al., 2012). Normalisation was set to 'false' since the standardisation was previously done during the alignment stage. Eleven harmonics were used for the lateral lamina and 12 for the terminal lamina, generating 44 and 48 harmonic coefficients respectively. Each set of harmonics had a cumulative power of 99% indicating that the number of harmonics generated was sufficient to describe the shape (Bonhomme et al., 2014). The harmonic coefficients of the terminal and lateral lamina were combined to form 92 harmonic coefficients together and this dataset was subsequently used for subsequent analyses as outlined below.

2.4.3 **Other morphological characters than leaf shape**

A morphological study of several additional vegetative characters ('Non-leaf shape dataset'; Table 2.2) was conducted to explore the potential taxonomic information that might

be obtained from vegetative characters other than leaf shape. For this purpose, a total of 52 specimens used in the phylogenetic study in Chapter 3 were investigated using 18 morphological characters (Table 2.2). Characters 8, 9, 14 (Table 2.2) were adopted from the Manual of Leaf Architecture (Leaf Architecture Working Group, 1999) which provides a comprehensive list of characters and character states for the study of leaf fossils. A calliper was used for measurements of quantitative characters. A dissecting microscope was used to examine the indumentum and any part of the leaf that needed detailed investigation, e.g. teeth on the leaflet margin.

Table 2.2. Non-leaf shape characters and character states used for morphology study of 52 Tetrastigma loheri *s.l. specimens.*

	Characters	Character states (type)
1	Maximum petiole length (cm)	(numerical)
2	Maximum petiolule length of the terminal leaflet (cm)	(numerical)
3	Maximum petiolule length of the lateral leaflet (cm)	(numerical)
4	Maximum internode length (cm)	(numerical)
5	Maximum tooth length	3 states: less than 1mm, 1mm, more than 1 mm (categorical).
6	Leaf margin	3 states: serrate, dentate, combination of serrate and dentate (categorical).
7	Maximum number of teeth on leaf margin per cm	(numerical)
8	Apical side tooth shape	6 states: different combinations of the following shapes: 1) concave and convex, 2) concave, convex, and straight, 3) convex and straight, 4) concave and straight, 5) straight, 6) convex (categorical).
9	Basal side tooth shape	6 states: different combinations of the following shapes: 1) concave and convex, 2) concave, convex, and straight, 3) convex and straight, 4) concave and straight, 5) straight, 6) convex (categorical).
10	Appearance of secondary veins	2 states: prominent, not prominent (categorical).
11	Maximum number of secondary veins on terminal leaflet	(numerical).
12	Maximum number of secondary veins on lateral leaflet	(numerical).
13	Secondary veins spacing	3 states: regular, irregular, mix of regular and irregular (categorical).
14	Inter-secondary veins thickness	4states: lacking, weak, strong, mix of weak and strong (categorical).
15	Tendril type	2 states: simple, forked (categorical)
16	Tendril indumentum (<1mm simple hairs)	3 states: absent, glabrescent, persistent (categorical).
17	Leaf indumentum (<1mm simple hairs)	2 states: absent, present (categorical).
18	Twig indumentum (<1mm simple hairs)	3 states: absent, glabrescent, persistent (categorical).

2.4.4 **Data analyses**

A total of five datasets were compiled for analysis: (1) the landmark dataset, (2) the leaf outline dataset, (3) a combined leaf shape dataset (landmark and leaf outline datasets), (4) the non-leaf shape dataset, and (5) a combined morphology dataset (leaf and non-leaf shape datasets). Data from 97 specimens were included in the first three datasets and data from 52 specimens in the latter two. The combined morphology dataset only included data from 52 specimens to avoid including missing data for samples for which only leaf shape

data was recorded. The combined leaf shape dataset contained a total of 152 characters. The combined morphology dataset contained 170 characters.

Different unsupervised clustering methods were used for each dataset to determine if the *T. loheri* s.l. complex is composed of groups of specimens that are morphologically distinct from each other. Unsupervised clustering methods aim to find groups of unknown composition and number in a dataset without prior information about the group membership of the specimens (Alashwal et al., 2019). The Principal Component Analysis (PCA), k-means clustering, and the model-based clustering method used in this study are only suitable for analyses using quantitative data. Because of this, these analyses were not used for the non-leaf shape and the combined morphology datasets, both of which are comprised of a mixture of quantitative and qualitative characters. Instead, these two datasets were analysed only using Principal Coordinate Analysis (PCoA), which can handle mixed data.

PCA is a multivariate analysis method that aims to generate a low dimensional set of variables from a high dimensional set of variables while maximizing the variance (Abdi & Williams, 2010). A linear combination of multiple variables is summarized into principal components which are used to plot a pattern of similarity of the data in a reduced number (usually two or three) of dimensions (Abdi & Williams, 2010). PCA can then be used as a graphical means of determining the presence of distinct clusters of specimens in morphometric space (Everitt & Wiley, 2011). For the landmark dataset, the Procrustes coordinates were converted into a covariance matrix that was used for PCA in MorphoJ. For the leaf outline dataset, PCA used the harmonic coefficients in R Studio using the *pca* function of Momocs. For the combined leaf shape dataset, the function *scale* was used to standardise the data before conducting a PCA using the function *prcomp* in R Studio.

PCoA is a multi-dimensional reductional method similar to PCA but uses a dissimilarity matrix as an input. Dissimilarity matrices of the non-leaf shape and combined morphology datasets were generated using Gower's distance which can handle a mixture of qualitative and quantitative data and allows for standardisation of the data (Gower, 1971). This was done in R studio with the function *daisy* with the metric type set to 'gower' using the statistical package *cluster* (Kaufman & Rousseeuw, 2009). The PCoA was performed in R studio using the function *cmdscale.*

K-means clustering aims to find clusters of specimens that meet the condition of having a minimum amount of variation within them and a maximum amount of variation among them using a k-means algorithm (MacQueen, 1967). The k-means algorithm was employed using the *kmeans* function in R Studio and the result was plotted using the function

fviz_cluster in the statistical package *factoextra* (Kassambara & Mundt, 2016). The optimal k for this analysis was identified using the gap statistic method, which compares the within-cluster variation for a series of k's and computes their expected values under the null distribution of the data (Tibshirani et al., 2001). The first highest gap statistic value indicates the best estimate of k. Data with well-separated clusters follow a decreasing trend from the highest gap statistic value (monotone; Tibshirani et al., 2001). A non-monotone behavior is characterised by the presence of several maxima and indicates the presence of overlapping subclusters within well-defined clusters (Tibshirani et al., 2001). Using the statistical package *factoextra,* the gap statistic method was employed in R Studio by using the function *fviz_nbclust* with 'gap_stat' as the method selection. The number of Monte Carlo bootstrap samples was set to 500 and the analysis was run for k up to 30, a range that is expected not to exceed the highest likely number of putative species.

Model-based clustering follows a finite mixture model which assumes the presence of underlying clusters in the data (Stahl & Sallis, 2012). It uses different statistical models to describe the clusters and finds the best way to optimise the fit between these models and the observed data (Gan et al., 2007). The best model is indicated by a high Bayesian Information Criterion (BIC) score (Fraley & Raftery, 1999). Model-based clustering was performed in R using the *mclust* statistical package (Fraley & Raftery, 1999) and the function *Mclust.* The result was plotted using the *fviz_mclust* function in the *factoextra* package.

2.5 RESULTS

2.5.1 Principal Component and Coordinate Analyses

The first principal component (PC1) of the PCA of the landmark dataset explains 43.84% of the variation in the leaf shapes while the second principal component (PC2) explains 16.41 % of the variation (60.25% combined). The PC1 and PC2 of the PCA of the leaf outline dataset explain 55.3% and 24.2% of the shape variation respectively and 79.5% combined. The PC1 and PC2 of the PCA of the combined leaf shape dataset explain 14.1% and 12.6% of the shape variation respectively (26.7% combined).

The PC1 and PC2 of the PCoA of the non-leaf shape data represent 12% and 8.9% of variation respectively with a total of 20.9%. The PC1 and PC2 of the PCoA of the combined morphology dataset represent the 12.7% and 11.1% of the variation with a 23.8% total variation for the two components.

The PCA ordination plots of the landmark, leaf outline, and combined leaf shape datasets (Figures 2.2, 2.3, and 2.4) and the PCoA ordination plots of the non-leaf shape and combined morphology datasets (Figures 2.5 and 2.6) do not show morphologically distinct clusters of specimens nor a clustering of specimens that were collected from the same province or from the same region (i.e. Luzon, Visayas, Mindanao region) in the Philippines (not shown).

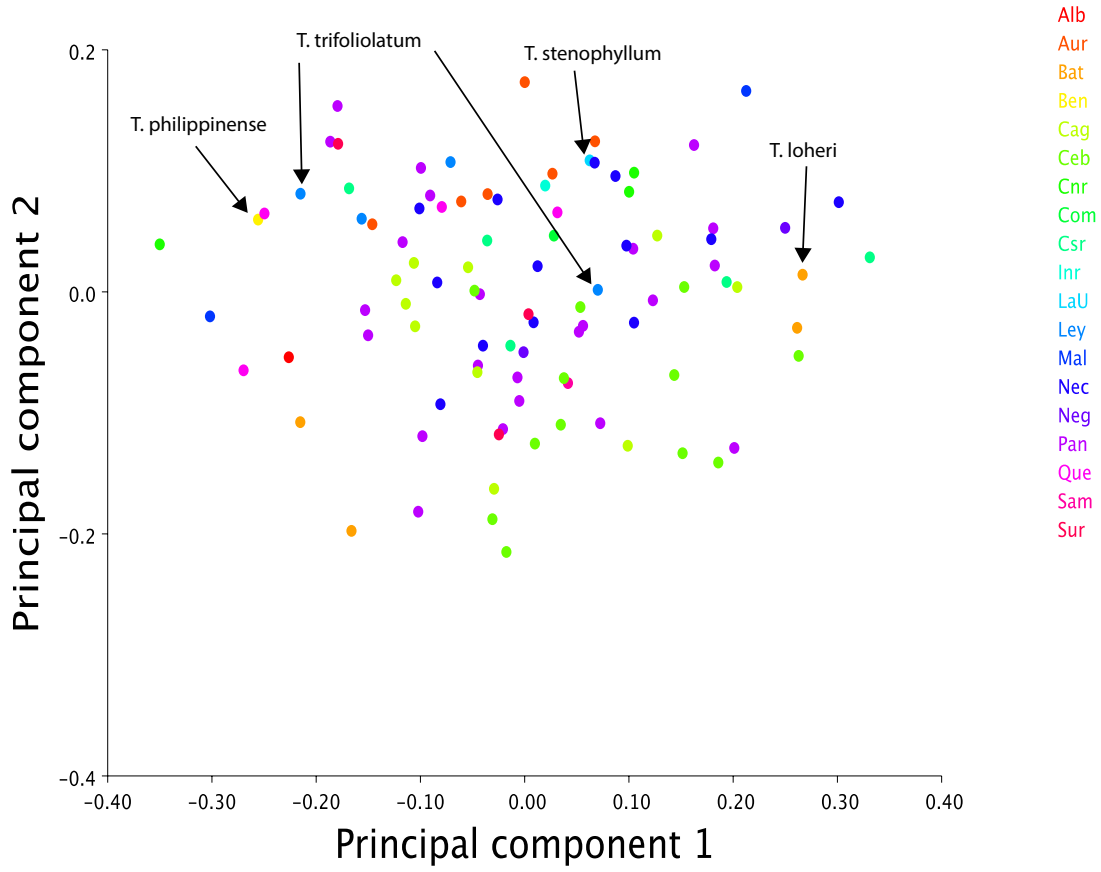


Figure 2.2. Ordination plot from a Principal Component Analysis of the landmark leaf shape data of 97 individuals of Tetrastigma loheri *s.l. using the first two principal components. Specimens collected from type localities of species are indicated. Legend refers to the Philippine province where the specimens were collected: Alb- Albay, Aur- Aurora, Bat- Bataan, Ben-Benguet, Cag-Cagayan, Ceb-Cebu, Cnr-Camarines Norte, Com- Compostela Valley, Csr-Camarines Sur, Inr- Ilocos Norte, LaU- La Union, Ley-Leyte, Mal- Albay (Mt. Malinao), Nec- Nueva Ecija, Neg-Negros Occidental, Pan-Antique & Iloilo, Que-Quezon, Sam-Samar, Sur-Surigao.*

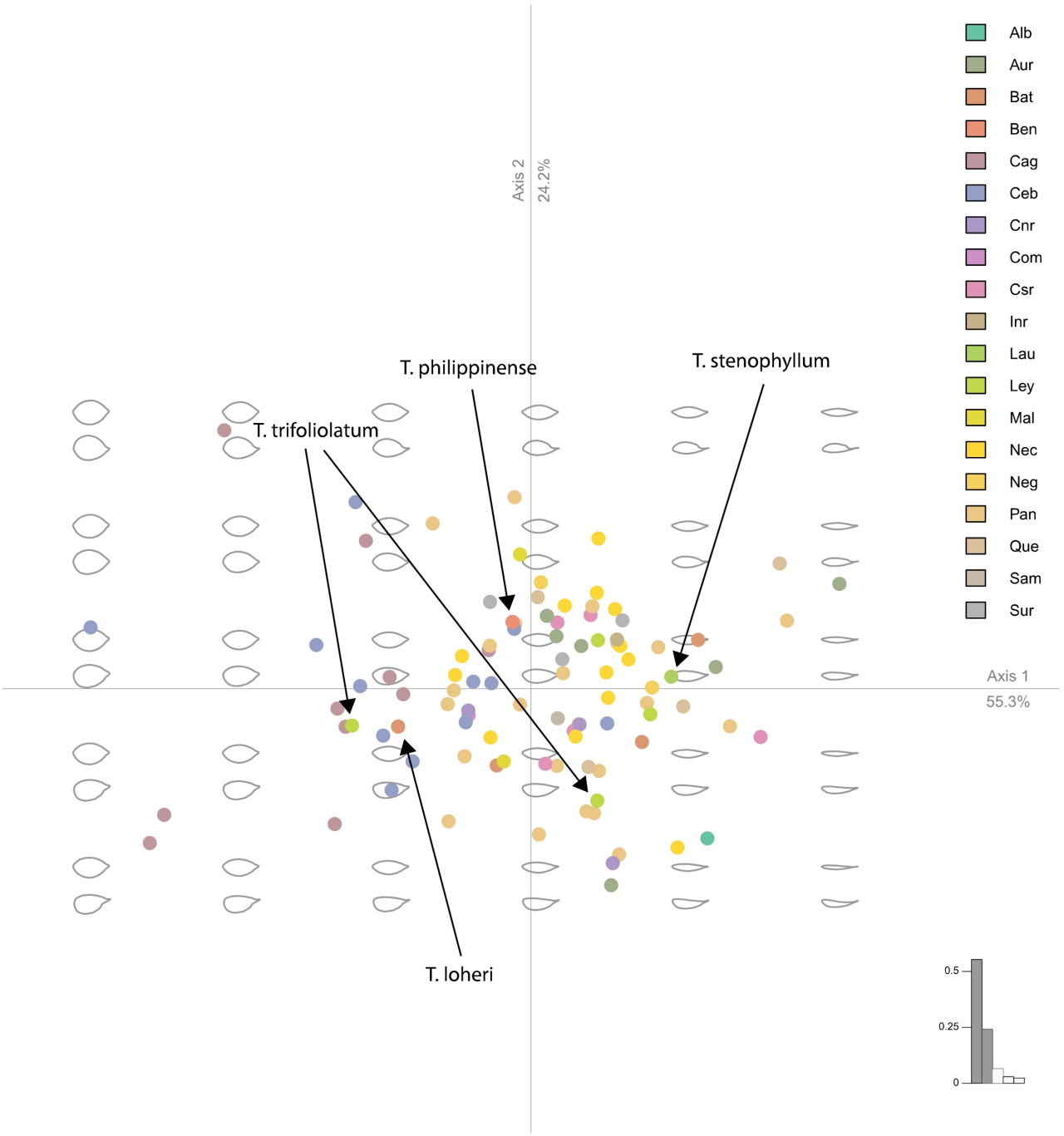


Figure 2.3. Ordination plot from a Principal Component Analysis of the leaf outline data of 97 individuals of Tetrastigma loheri *s.l. using the first two principal components. Specimens collected from type localities of species are indicated. The paired ellipsoid shapes refer to the leaflet lamina shapes in which the top corresponds to the terminal leaflet and the bottom to the lateral leaflet. Legend refers to the Philippine province where the specimens were collected: Alb- Albay, Aur- Aurora, Bat-Bataan, Ben-Benguet, Cag-Cagayan, Ceb-Cebu, Cnr-Camarines Norte, Com- Compostela Valley, Csr-Camarines Sur, Inr- Ilocos Norte, LaU-La Union, Ley-Leyte, Mal- Albay (Mt. Malinao), Nec- Nueva Ecija, Neg-Negros Occidental, Pan-Antique & Iloilo, Que-Quezon, Sam-Samar, Sur-Surigao.*

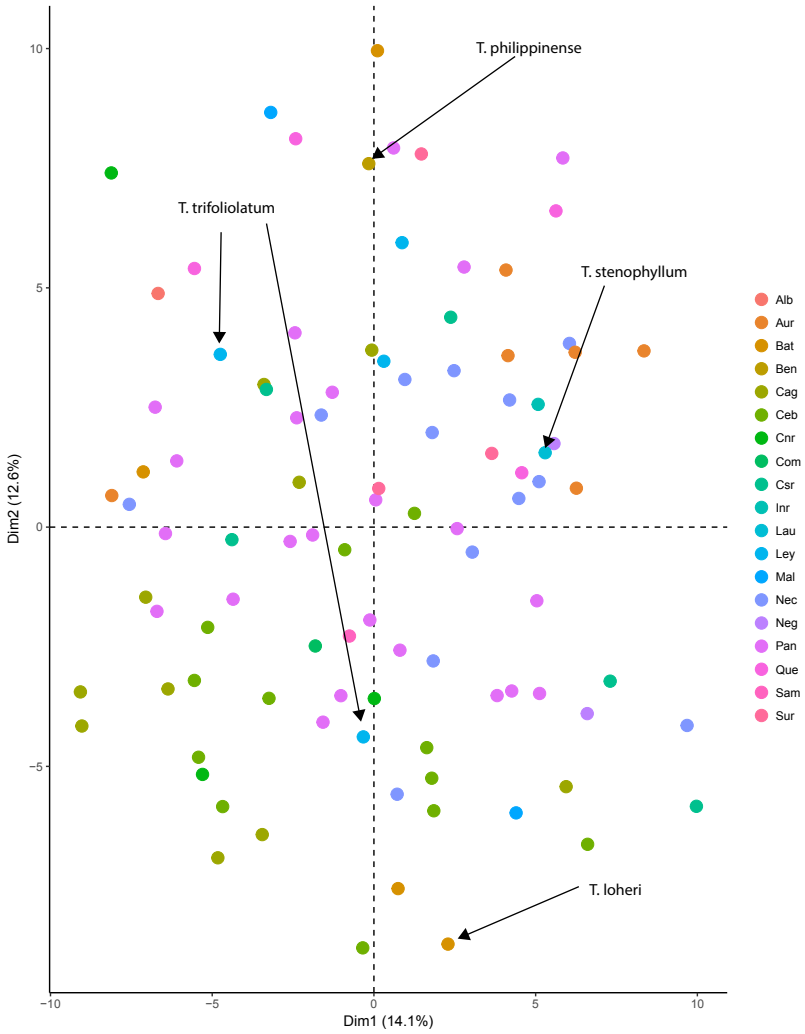


Figure 2.4. Ordination plot from a Principal Component Analysis of the combined leaf shape data (landmark and leaf outline) of 97 individuals of Tetrastigma loheri *s.l. showing the first two principal components. Specimens collected from type localities of species are indicated. Legend refers to the Philippine province where the specimens were collected: Alb- Albay, Aur- Aurora, Bat-Bataan, Ben-Benguet, Cag-Cagayan, Ceb-Cebu, Cnr-Camarines Norte, Com- Compostela Valley, Csr-Camarines Sur, Inr- Ilocos Norte, LaU- La Union, Ley-Leyte, Mal- Albay (Mt. Malinao), Nec- Nueva Ecija, Neg-Negros Occidental, Pan-Antique & Iloilo, Que-Quezon, Sam-Samar, Sur-Surigao.*

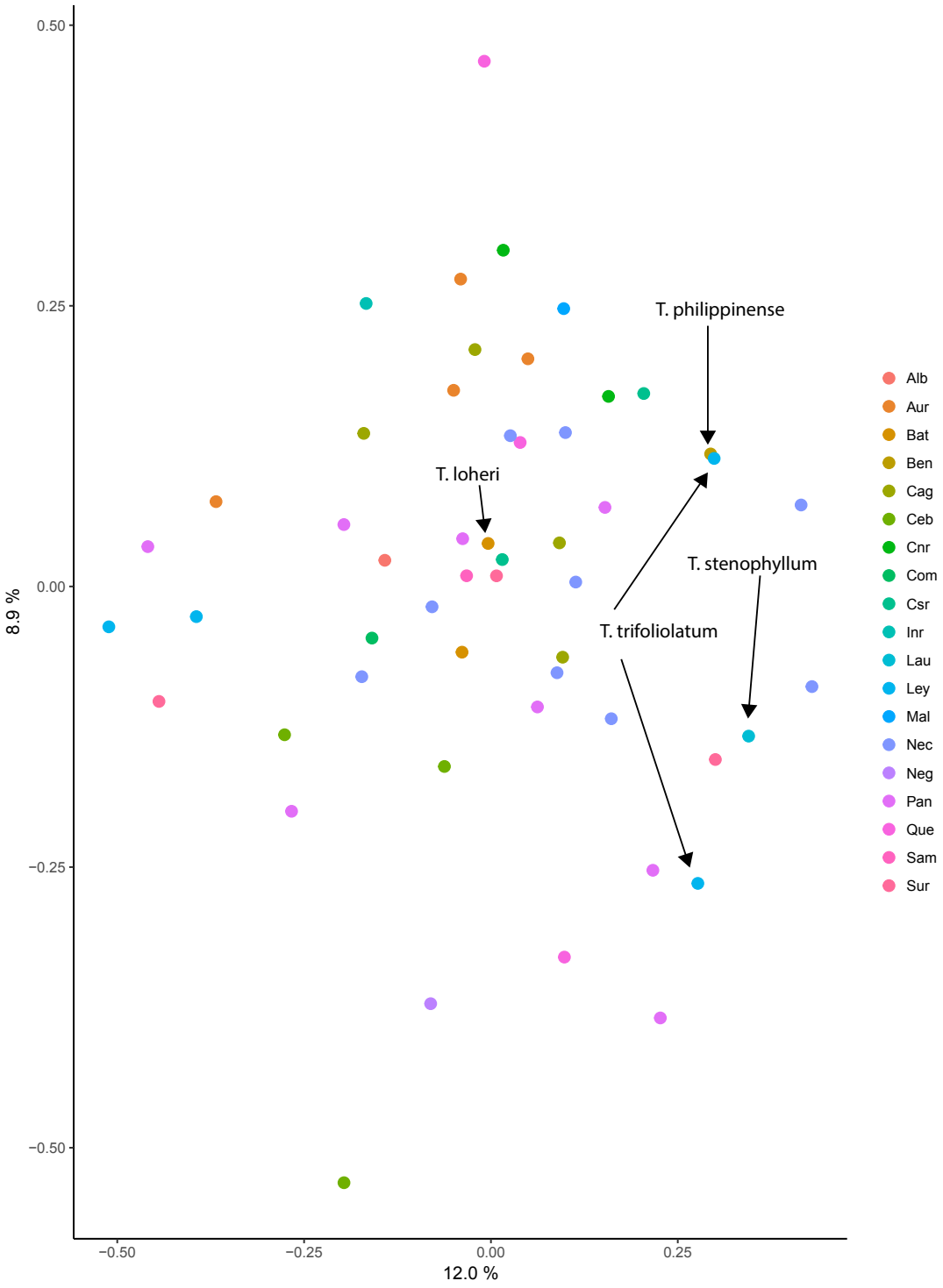


Figure 2.5. Ordination plot from a Principal Coordinate Analysis of the non-leaf shape data of 52 individuals of Tetrastigma loheri *s.l. using the first two principal components. Specimens collected from type localities of species are indicated. Legend refers to the geographic origin of the specimens: Alb- Albay, Aur- Aurora, Bat-Bataan, Ben-Benguet, Cag-Cagayan, Ceb-Cebu, Cnr-Camarines Norte, Com- Compostela Valley, Csr-Camarines Sur, Inr- Ilocos Norte, LaU- La Union, Ley-Leyte, Mal- Albay (Mt. Malinao), Nec- Nueva Ecija, Neg-Negros Occidental, Pan-Antique & Iloilo, Que-Quezon, Sam-Samar, Sur-Surigao.*

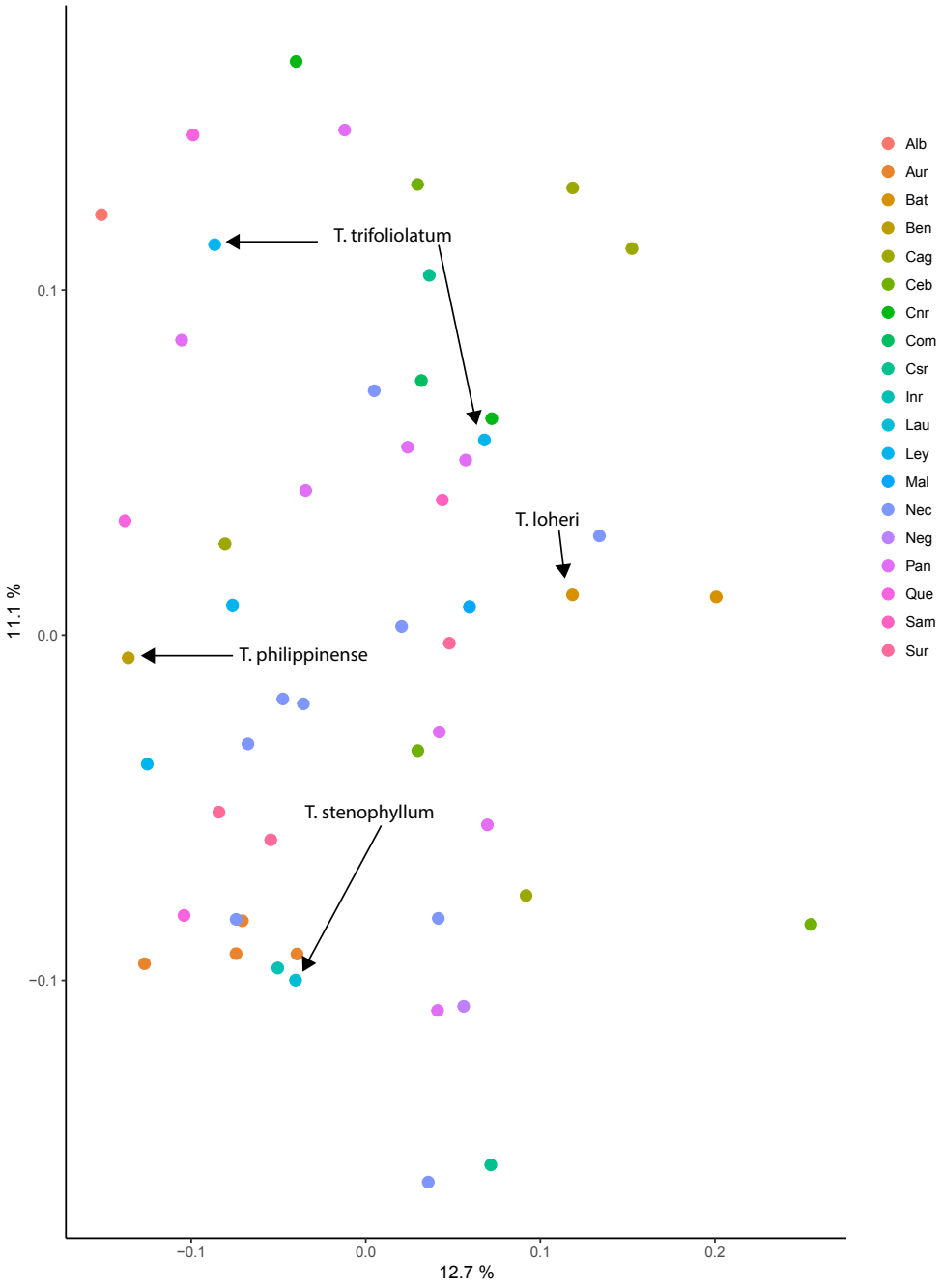


Figure 2.6. Ordination plot from a Principal Coordinate Analysis of the combined morphology data (leaf shape and non-leaf shape) of 52 individuals of Tetrastigma loheri *s.l. using the first two principal components. Specimens collected from type localities of species are indicated. Legend refers to the geographic origin of the specimens: Alb- Albay, Aur-Aurora, Bat-Bataan, Ben-Benguet, Cag-Cagayan, Ceb-Cebu, Cnr-Camarines Norte, Com-Compostela Valley, Csr-Camarines Sur, Inr- Ilocos Norte, LaU- La Union, Ley-Leyte, Mal-Albay (Mt. Malinao), Nec- Nueva Ecija, Neg-Negros Occidental, Pan-Antique & Iloilo, Que-Quezon, Sam-Samar, Sur-Surigao.*

2.5.2 **K-means and model-based clustering**

2.5.2.1 Landmark dataset

Both the k-means and model-based clustering failed to find more than one cluster of specimens in the landmark dataset (ordination plots not shown). The results of the gap statistic analysis as part of the k-means analysis show a non-monotone pattern with the first maximum gap statistic found at k=1 (Figure 2.7). A generally increasing trend towards k=30 is evident after k=2, with several maxima. The results of the model-based clustering show that the best model fitted for the data is the ellipsoidal multivariate normal model with k=1, which has the highest BIC of 44186.04 (Figure 2.8).

2.5.2.2 Leaf outline dataset

A non-monotone trend is evident from the results of the gap statistic analysis where the first maximum was obtained at k=1 (Figure 2.9). A generally increasing trend towards k=30 with several maxima after k=3 was found. The ordination plot resulting from the k-means analysis is not shown.

In contrast to the results of the k-means analysis, the model-based clustering method shows that the best k is five, based on the ellipsoidal, equal volume, shape, and orientation model which has the highest BIC value of 5480.49 (Figure 2.10). The associated ordination plot shows one large cluster that consists of 93 individuals (and includes the specimens collected at type localities) and four small clusters which are each composed of a single individual (Figure 2.11). The individuals in these four small clusters are from Cagayan and Panay.

2.5.2.3 Combined leaf shape dataset

The first highest gap statistic for the k-means analysis is found at k=1. An increasing trend towards k=30 after the lowest point at k=3 is evident and shows several maxima (Figure 2.12).

The model-based clustering (Figure 2.13) indicates that the optimal k is four based on the spherical, varying volume model with highest BIC value of -40402.05. The associated ordination plot shows that these four clusters are overlapping (Figure 2.14). Each cluster is comprised of specimens from different provinces. The *T. loheri* specimen from the type locality of this species is found in the largest cluster whereas those of *T. stenophyllum*, *T. trifoliolatum* and *T. philippinense* are in the three smaller clusters.

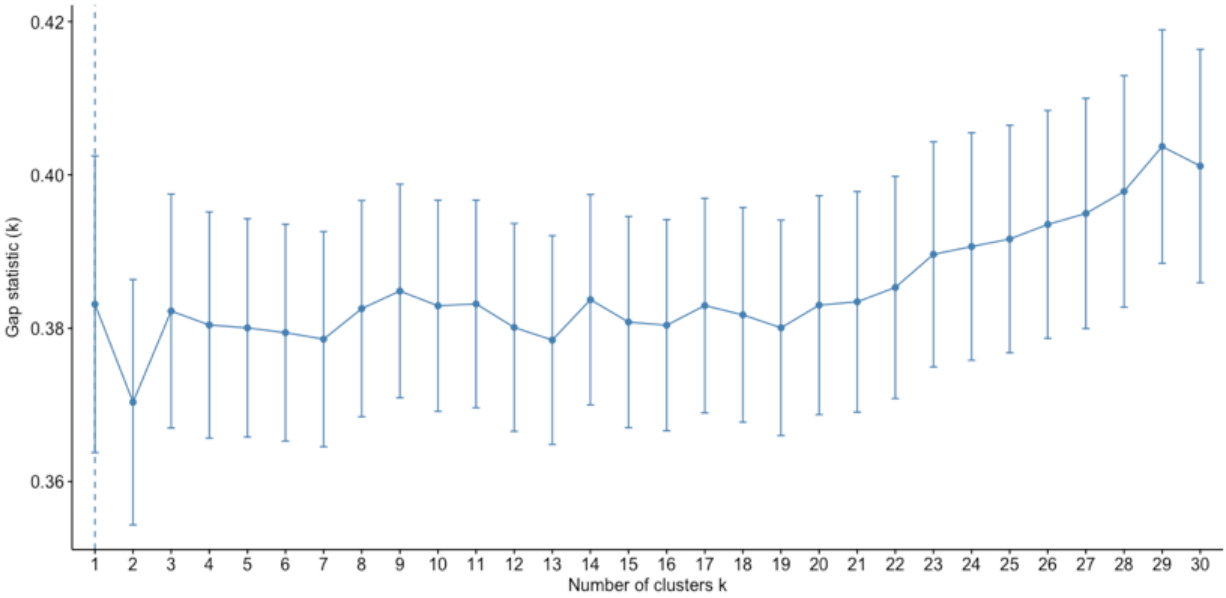


Figure 2.7. Results of the gap statistic analysis (k-means analysis) for the landmark dataset for Tetrastigma loheri *s.l.*

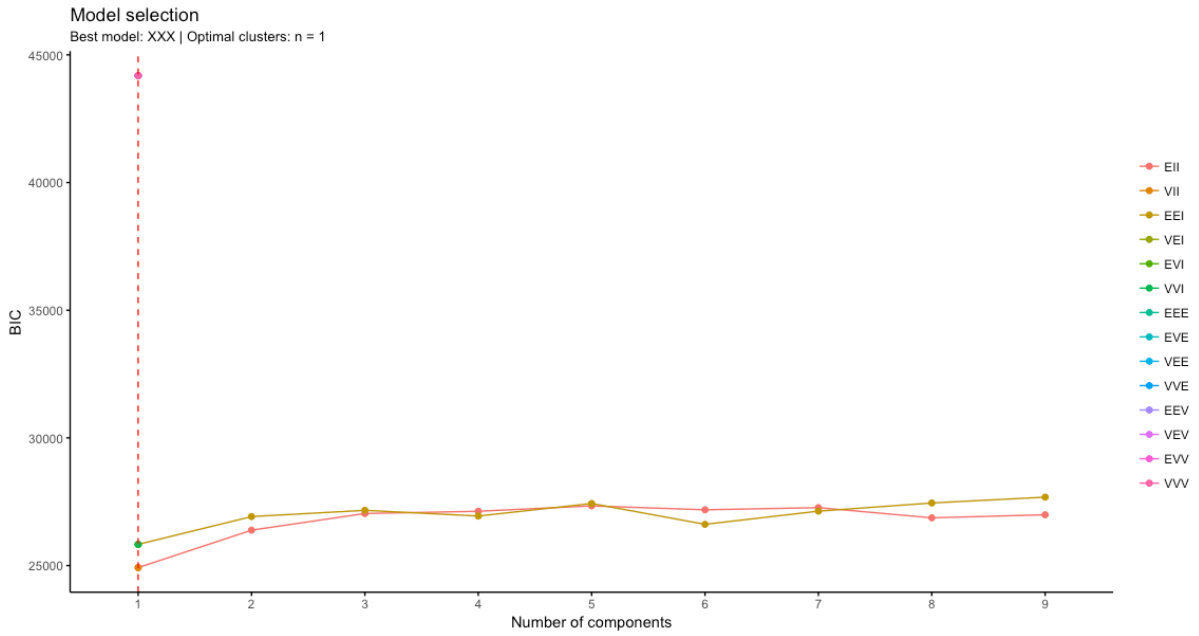


Figure 2.8. Optimal k identification for model-based clustering for the landmark dataset for Tetrastigma loheri *s.l. BIC: Bayesian information criterion. The legend on the right shows the different models included in the Mclust statistical package.*

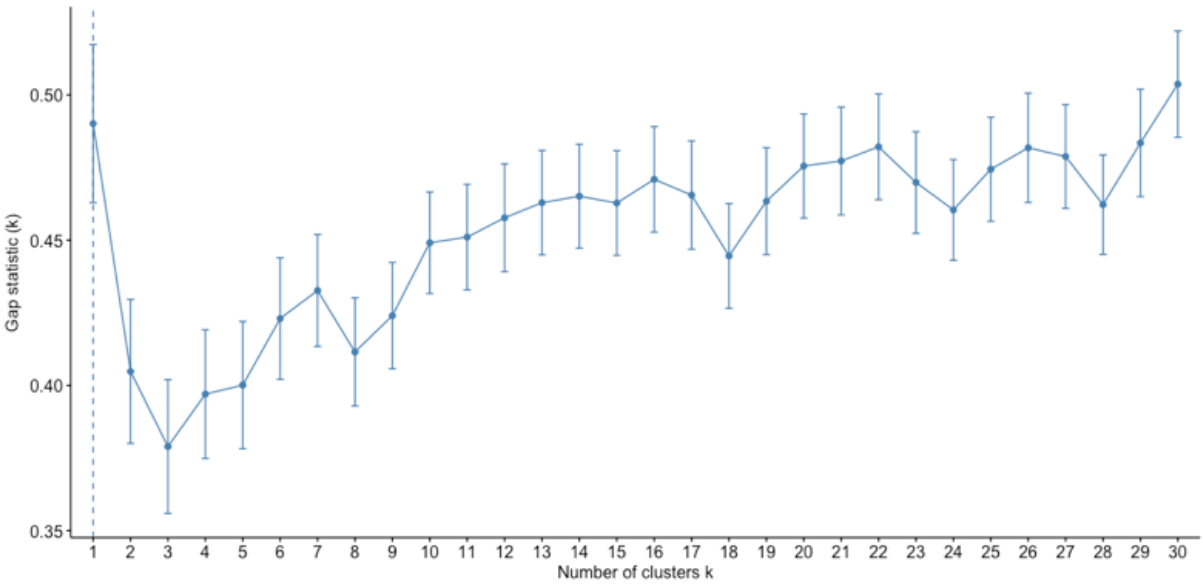


Figure 2.9. Results of the gap statistic analysis (k-means analysis) for the leaf outline dataset for Tetrastigma loheri *s.l.*

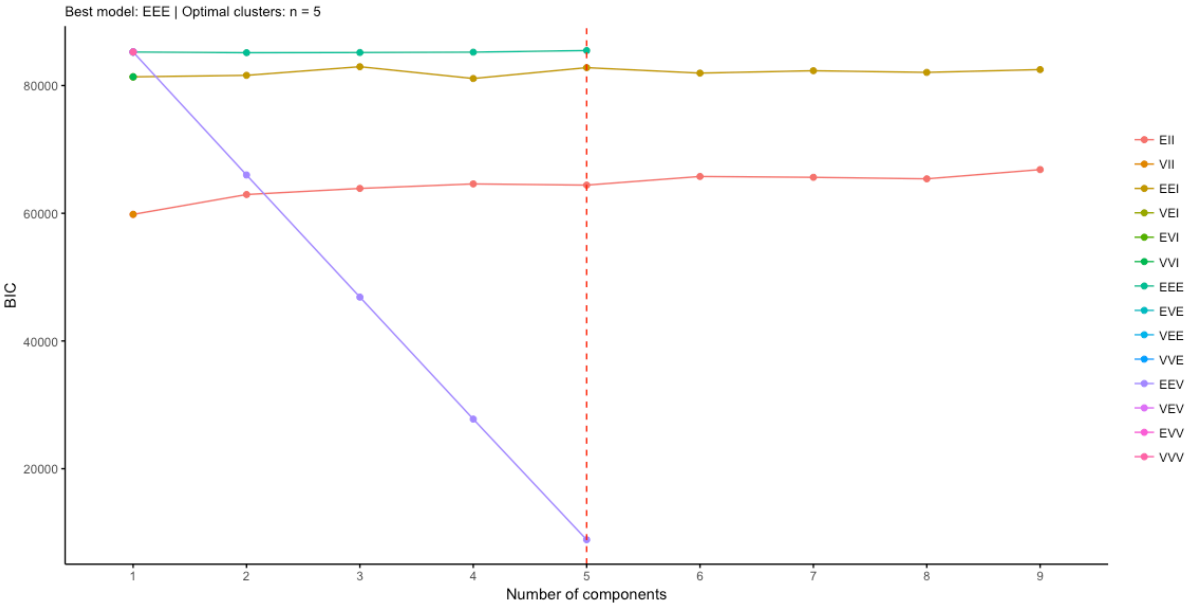


Figure 2.10. Optimal k identification for model-based clustering for the outline dataset for Tetrastigma loheri *s.l. BIC: Bayesian information criterion. The legend on the right indicates the different models included in the Mclust statistical package.*

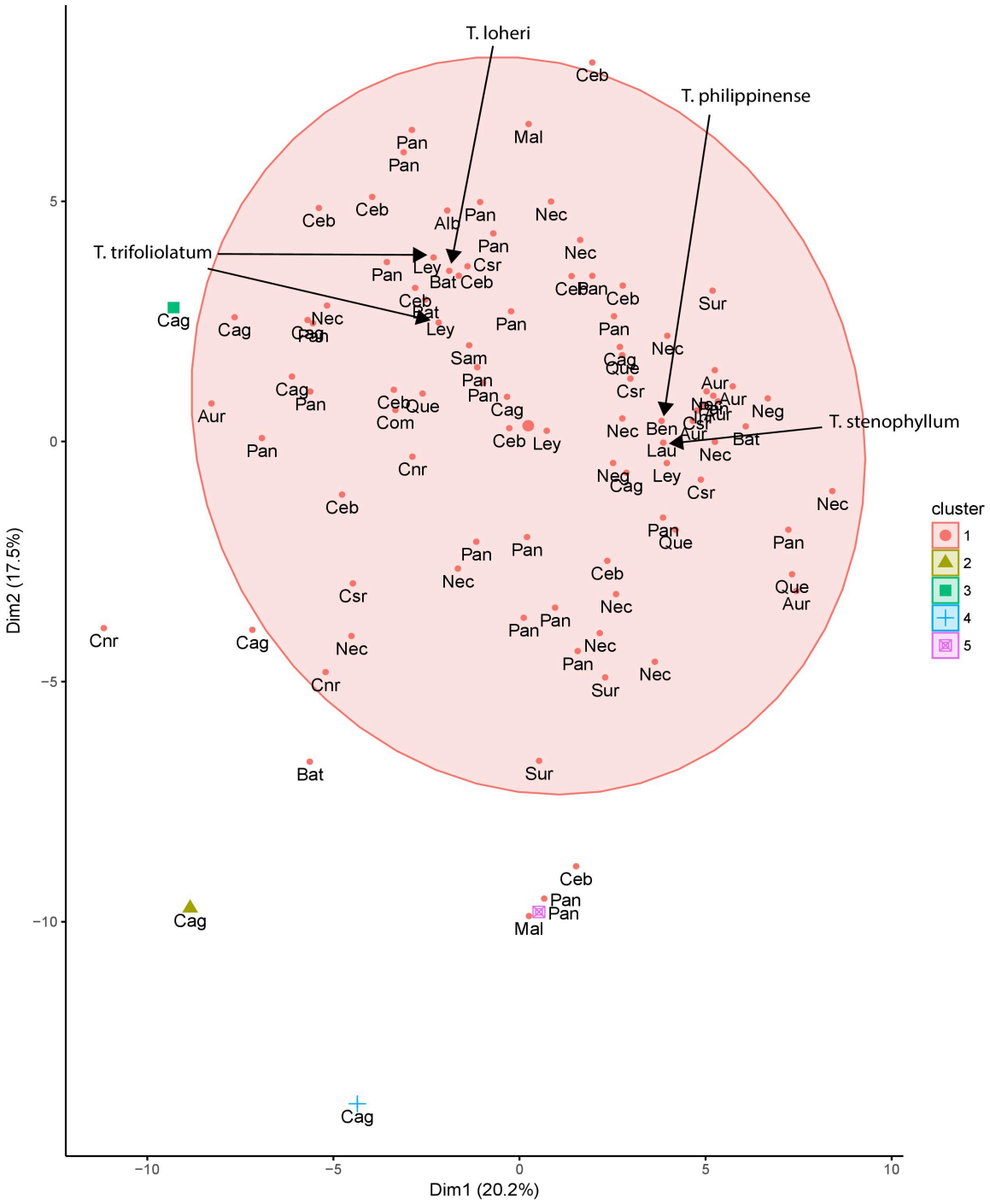


Figure 2.11. Ordination plot from a model-based clustering analysis of the leaf outline data of 97 individuals of Tetrastigma loheri *s.l. using the first two principal components. One cluster is represented by one ellipse at the 90% confidence level. The other four clusters are comprised of one individual each (ellipses not created). Specimens collected from type localities of species are indicated. Legend refers to the Philippine province where the specimens were collected: Alb- Albay, Aur- Aurora, Bat-Bataan, Ben-Benguet, Cag-Cagayan, Ceb-Cebu, Cnr-Camarines Norte, Com- Compostela Valley, Csr-Camarines Sur, Inr- Ilocos Norte, LaU- La Union, Ley-Leyte, Mal- Albay (Mt. Malinao), Nec- Nueva Ecija, Neg-Negros Occidental, Pan-Antique & Iloilo, Que-Quezon, Sam-Samar, Sur-Surigao.*

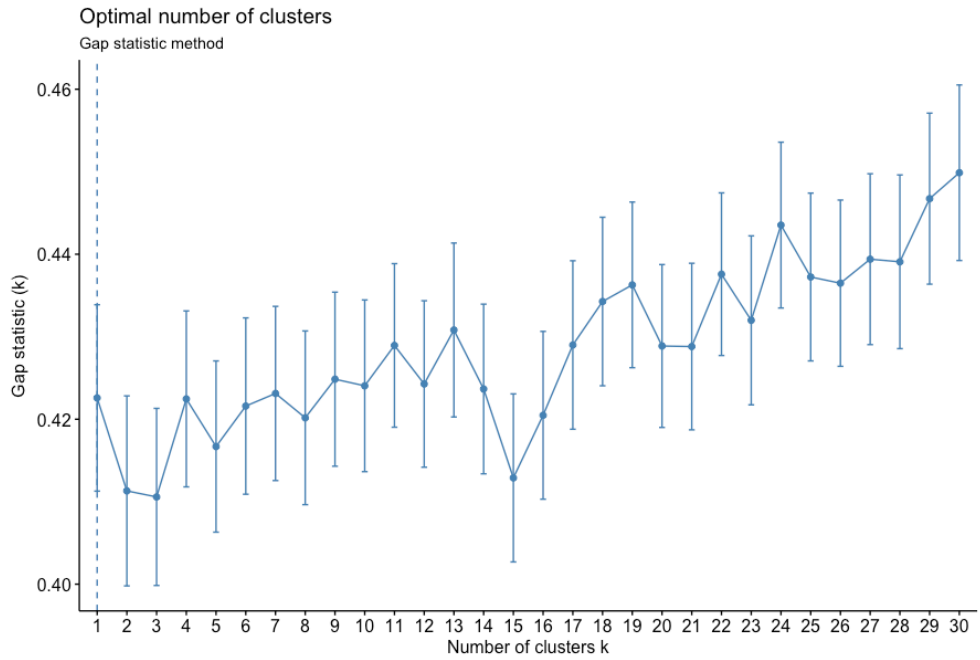


Figure 2.12. Results of the gap statistic analysis (k-means analysis) for combined landmark and outline dataset for Tetrastigma loheri *s.l.*

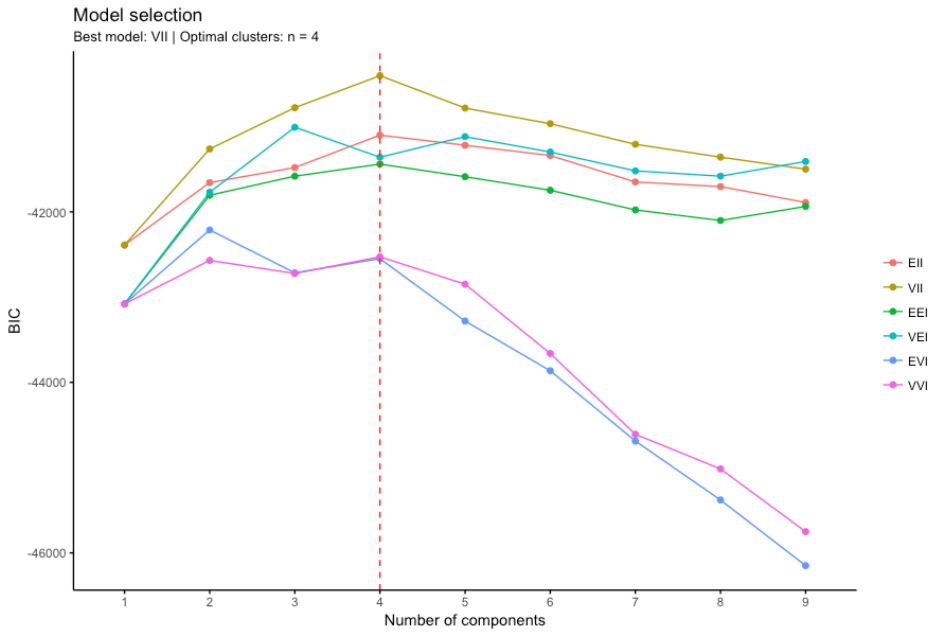


Figure 2.13. Optimal k identification for model-based clustering for the combined leaf shape dataset for Tetrastigma loheri *s.l. BIC: Bayesian information criterion. The legend on the right indicates the different models included in the Mclust statistical package.*

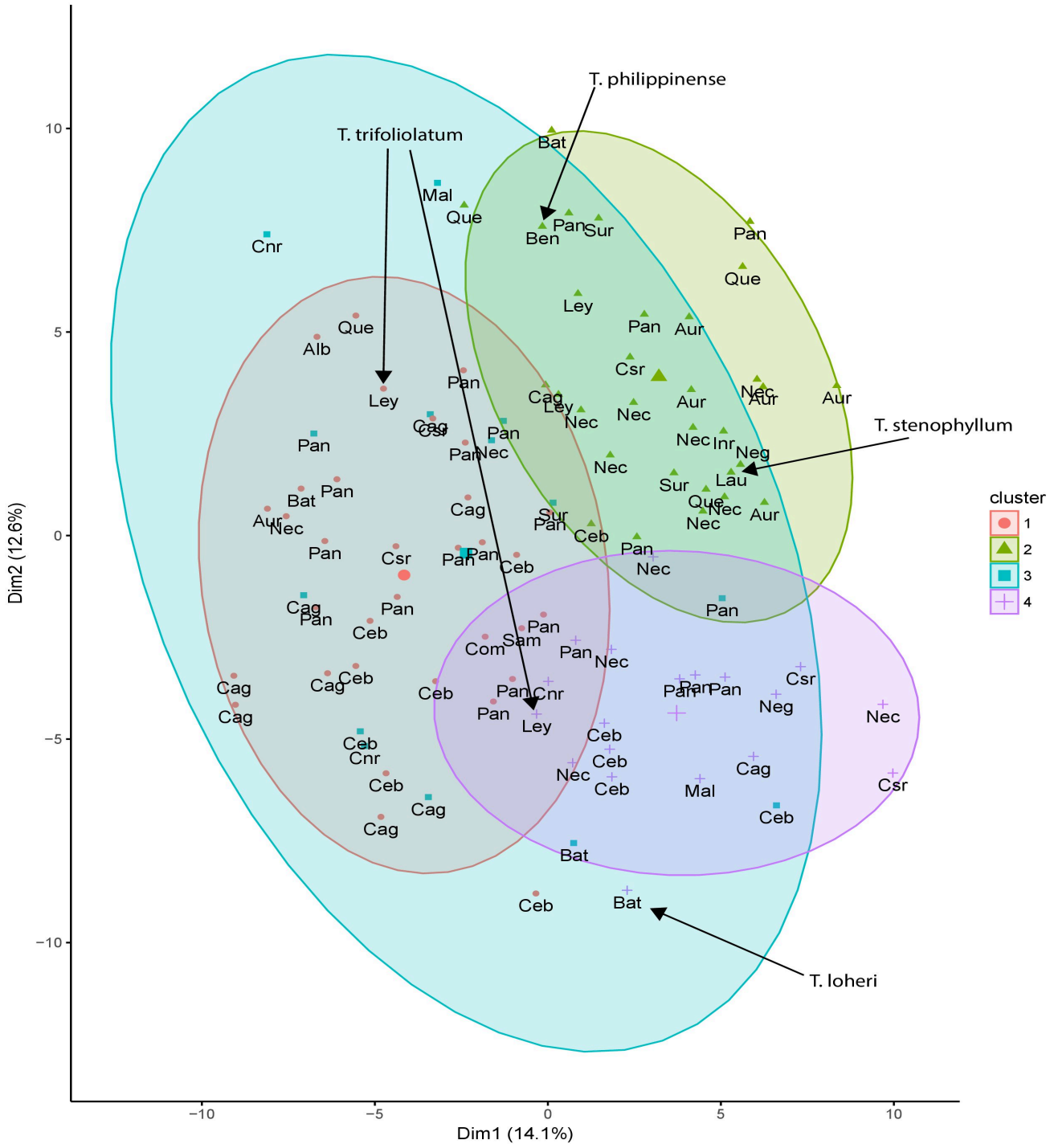


Figure 2.14. Ordination plot from the model-based clustering analysis of the combined leaf shape dataset (landmark and leaf outline) of 97 individuals of Tetrastigma loheri *s.l. using the first two principal components. The four clusters are represented by four ellipses of 90% confidence level. Specimens collected from type localities of species are indicated. Legend refers to the regional origin of the specimens: Alb- Albay, Aur- Aurora, Bat-Bataan, Ben-Benguet, Cag-Cagayan, Ceb-Cebu, Cnr-Camarines Norte, Com- Compostela Valley, Csr-Camarines Sur, Inr- Ilocos Norte, LaU- La Union, Ley-Leyte, Mal- Albay (Mt. Malinao), Nec- Nueva Ecija, Neg-Negros Occidental, Pan-Antique & Iloilo, Que-Quezon, Sam-Samar, Sur-Surigao.*

2.6 DISCUSSION

Tetrastigma loheri will be used in the research presented in Chapter 4 of this thesis for studying patterns of genetic diversity and connectivity among four remaining forest areas in Cebu. However, the results of previous research (Pelser et al., 2016b) indicate that this species potentially forms a species complex in the Philippines. Because of morphological similarities, this complex potentially also includes Philippine plants previously identified as *T. diepenhorstii, T. philippinense, T. stenophyllum,* and *T. trifoliolatum.* The present study aimed to determine if there is more than one species in *T. loheri* s.l. by conducting a morphological study of vegetative characters with the goal of finding groups of specimens that are morphologically distinct. Such groups could indicate the presence of more than one species in the complex under the unified species concept (De Queiroz, 2007) using morphological distinction as evidence of lineage separation.

2.6.1 Species delimitation of the *T. loheri* s.l. complex

Geometric morphometric methods have proved useful in resolving the taxonomic delimitation of morphologically diverse species complexes and other taxonomically complicated groups (e.g., Nery & Fiaschi, 2019). Here, two geometric morphometric approaches were used to study patterns of variation in leaf shape morphology within the *T. loheri* s.l. complex. In addition, a select number of other vegetative characters were studied.

The results of the majority of the morphometric analyses of the different datasets that were compiled indicate that the vegetative characters that were studied do not provide evidence for recognizing more than one species within the *T. loheri* s.l. complex in the Philippines. However, the results of the model-based clustering analyses of the leaf outline and the combined leaf shape datasets (Figures 2.11 and 2.14) reveal the presence of five and four clusters of specimens, respectively, therefore indicating instead that more than one species might exist within the complex. The clustering patterns obtained from these two data sets are, however, incongruent with each other in the number of clusters identified as well as in their delimitation (Figures 2.11 and 2.14). The collective results of this study therefore provide more support for the existence of one species in the complex than four or five.

If the complex is indeed only composed of one species, this means that it displays substantial morphological variation, particularly in its leaf shape. The PCA ordination plot of the leaf outline dataset, for example, shows large variation in the length-width ratio of both the lateral and terminal leaflets (Figure 2.3). Although the morphological extremes in this ordination plot are remarkably different in this regard, the plot shows a continuous gradient

of morphological variation in leaflet shape (Figure 2.3). The ordination results do not reveal obvious geographic patterns in the morphological variation observed. For example, specimens from the island of Panay are positioned throughout the morphometric space of the ordination plots (e.g., Figure 2.14). In addition, specimens from the same locality can be substantially different in their leaf morphology. This is for instance evident from the different positions in the ordination plots of the two specimens that were identified as *T. trifoliolatum* and that were collected at the same locality (i.e. Barangay Camono-an in Dagami; e.g., Figures 2.3, 2.4, and 2.14). However, the examination of individual characters (results not shown) revealed regional differences in tendril morphology. Whereas plants from the islands of Cebu, Leyte, Samar, and Negros in the central Visayas region have forked tendrils (e.g., *Barcelona 4007*, CANU; Nickrent et al., 2006 onwards: DOL103566), those from other parts of the Philippines exclusively display simple tendrils. Morphological differences that are associated with habitat, elevation, or other ecological factors were not observed (results not shown). Overall, under the assumption that the *T. loheri* s.l. complex is only composed of a single species, it therefore seems most likely that the large morphological diversity in the vegetative characters that were studied is a result of phenotypic plasticity, although local adaptation cannot be entirely excluded.

2.6.2 The taxonomic status of *T. diepenhorstii*, *T. loheri, T. philippinense, T. stenophyllum*, and *T. trifoliolatum*

The results of this study do not provide convincing evidence that the *T. loheri* s.l. complex is composed of more than one species in the Philippines. Under the assumption that this is correct, this might have implications for the taxonomic status and/or the delimitation of *T. loheri*, *T. philippinense, T. stenophyllum,* and *T. trifoliolatum*. As circumscribed in their protologues (Gagnepain, 1910; Merrill, 1912, 1914, 1916), these taxa are morphologically similar to each other, indicating that they are potentially closely related. In fact, the existence of specimens that cannot conclusively be assigned to a particular species in this assemblage suggests that these taxa are even less distinct than previously assumed. For example, some specimens that were included in this study display character states that are diagnostic for one species in combination with character states that are diagnostic for a different species. For instance, *Callado 390* is characterised by a combination of character states such as 15 cm terminal lamina length, 13 cm lateral lamina length, and 3.6 cm terminal petiole length, each of which is a diagnostic leaf character state for both *T. philippinense* and *T. trifoliolatum, T. philippinense,* and *T. loheri* respectively. In addition, some specimens have character states

that are intermediate between those presented as diagnostic for different species, blurring the distinction between them. For example, *Barcelona 3826 with Pelser* has a combination of character states such as 8.9 cm petiole length, 5.2 cm terminal petiolule length, and 1.8 cm lateral petiolule length, which are all found between the diagnostic leaf characters states for *T. loheri*, *T. philippinense, T. stenophyllum,* and *T. trifoliolatum*. These patterns suggest that *T. loheri*, *T. philippinense, T. stenophyllum,* and *T. trifoliolatum* might even be conspecific. Although type specimens of these names could not be included in this study, this hypothesis was explored by including specimens that were collected from the type localities. Because the morphological features of these specimens conform to those of the four species as outlined in their protologues, they most likely indeed represent *T. loheri*, *T. philippinense, T. stenophyllum,* and *T. trifoliolatum*. For example, the position of the specimen from the type locality of *T. stenophyllum* in the PCA ordination plot of the leaf outline dataset (Figure 2.3) shows the narrow leaflets that are diagnostic for this species (Merrill, 1916).

The results of all analyses of all datasets that were used in this study place the specimens from the type localities of the four species among the other *T. loheri* s.l. specimens in the ordination plots (e.g., Figure 2.11). Under the assumption that the vegetative characters that were studied are suitable for studying the species delimitation of the *T. loheri* s.l. complex, this suggests that *T. loheri*, *T. philippinense, T. stenophyllum,* and *T. trifoliolatum* are conspecific and their names synonymous. As per the ICN (Turland et al., 2018), the name *T. loheri* Gagnep. has priority and should therefore be used to refer to this species.

The taxonomic status of *T. diepenhorstii* and the taxonomic identity of plants identified as *T. loheri* and *T. trifoliolatum* outside of the Philippines remain unclear. Specimens of the latter two were not available for this study. Although they might belong to *T. loheri*, it is also possible that they are representatives of a different, yet morphologically similar species that was not included in this study. *Tetrastigma diepenhorstii* is potentially such a species. Other than a single specimen (*Wen 8261*, US), which was not available for this study but was resolved as nested within a clade of *T. loheri* s.l. specimens in a molecular phylogenetic study (Pelser et al., 2016b), *T. diepenhorstii* has not been recorded for the Philippines (Pelser et al. 2011 onwards). If this specimen is correctly identified, its phylogenetic position would suggest that *T. loheri* and *T. diepenhorstii* are synonymous. By extension, this taxonomic view is embraced by Latiff (2001) and Zakaria et al. (2016, 2017), who concluded that *T. diepenhorstii* and *T. trifoliolatum* are synonyms. If correct, the name *T. diepenhorstii* (Miq.) Latiff has priority over *T. loheri* Gagnep. and should therefore be used for the plants that constitute the *T. loheri* s.l. complex as referred to in this thesis.

2.6.3 **Caveats and further investigation**

Although the results of the present study suggest that leaf shape data and other vegetative characters do not provide evidence for recognizing multiple species within the *T. loheri* s.l. complex, this hypothesis cannot be rejected conclusively.

Because of a lack of specimens with reproductive structures, patterns of morphological variation in inflorescence, flower and fruit morphology were not studied. It is therefore possible that the *T. loheri* s.l. complex is composed of species that are cryptic in their vegetative morphology, but distinct in reproductive characters. Although incomplete description of reproductive morphology is evident in the protologues or other taxonomic treatments of *T. diepenhorstii*, *T. loheri*, *T. philippinense, T. stenophyllum,* and *T. trifoliolatum* (Miquel, 1861; Gagnepain, 1910; Merrill, 1912, 1914, 1916; Zakaria et al., 2016, 2017), a few differences among species are found using available information. For example, inflorescence is pubescent in *T. diepenhorstii, T. stenophyllum, T. trifoliolatum*, and *T. philippinense* while it is glabrous in *T. loheri*; petals are pubescent in *T. stenophyllum* and *T. philippinense* while these are glabrous in *T. loheri* (no explicit mention for *T. diepenhorstii* and *T. trifoliolatum*); ovary is pubescent is *T. stenophyllum* while it is glabrous in *T. philippinense* (no explicit mention for *T diepenhorstii, T. loheri*, and *T. trifioliolatum*). These few differences indicate the possibility that the species comprising *T. loheri* complex might be distinguishable if reproductive characters are further investigated.

Furthermore, it is possible that the *T. loheri* s.l. complex is composed of species that are morphologically different in some of the vegetative characters that were studied, but that these differences are too subtle to be identified in the unsupervised clustering analyses that were used. In Chapter 3, a supervised method (Random Forest analysis) is therefore used to determine if putative species identified using phylogeny-based species delimitation models are distinct in their vegetative morphology.

CHAPTER 3: Species delimitation modelling does not provide evidence for the existence of more than one species within the *Tetrastigma loheri* s.l. species complex in the Philippines

3.1 ABSTRACT

Tetrastigma loheri (Vitaceae) is a vine species that is common in forests in Cebu, a large island in the Philippines. It is studied in this thesis to provide the first information about the extent of the effects of habitat fragmentation on patterns of genetic diversity and connectivity for four of the few remaining forested areas on the island. However, there are indications that *T. loheri* is part of a species complex in the Philippines that potentially also contains Philippine plants identified as *T. philippinense, T. diepenhorstii, T. stenophyllum,* and *T. trifoliolatum* (i.e. *T. loheri* sensu lato). This uncertainty about its species delimitation can complicate conservation genetic studies using this species (Chapter 4). In Chapter 2, I concluded that unsupervised morphometric analyses of leaf shape data and several other vegetative morphological characters do not support the existence of more than one species within *T. loheri* s.l. in the Philippines. In this Chapter, I continued testing this hypothesis by using Generalized Mixed Yule Coalescent (GMYC) and Poisson Tree Process (PTP) species delimitation models to identify clades within DNA sequence phylogenies of *T. loheri* s.l. that might constitute species within this complex. Although these methods identified several putative species, subsequent Random Forest analyses of the vegetative characters that were studied in Chapter 2 did not result in the identification of characters that can be used to discriminate them morphologically. Therefore, under a unified species concept that uses monophyly and morphological distinction as criteria for species recognition, I conclude that the currently available data do not support recognizing more than one species in the *T. loheri* s.l. complex. If correct, this implies that *T. philippinense, T. stenophyllum,* and *T. trifoliolatum* are possibly not taxonomically distinct from *T. loheri*, because specimens from the type localities of these species were found nested among other *T. loheri* s.l. specimens in the DNA sequence phylogenies. It is also possible that *T. diepenhorstii* is synonymous.

3.2 INTRODUCTION

It is estimated that only 0.2% of the original forest cover of the island of Cebu in the Philippines remains (Seidenschwarz, 2013). However, its few and small remaining forested areas still support endemic and threatened biodiversity (Paguntalan et al., 2015) which likely has experienced population decline and isolation due to habitat destruction, degradation and fragmentation (Young et al., 1996). Loss of genetic diversity is predicted in such small and isolated areas, resulting in the reduction of the evolutionary potential of populations of species that rely on these habitats, and, therefore of their chances of long-term persistence (Ellstrand & Elam, 1993; Jamieson et al., 2008). Data about patterns of genetic diversity and genetic connectivity can inform conservation management aimed at reducing genetic diversity loss in isolated habitats (Jamieson et al., 2008), however, these data are currently lacking for the remaining forested areas of Cebu. The research presented in Chapter 4 aims to contribute to addressing this knowledge gap by documenting patterns of genetic diversity and connectivity of *Tetrastigma loheri* Gagnep. (Vitaceae) in four of Cebu's forested areas. *Tetrastigma loheri* is a suitable species for this, because it is a commonly encountered forest vine species in all four selected forest areas and can therefore provide data about genetic connectivity among all four areas and a sufficient number of individuals per area to yield statistically well-supported genetic patterns. However, as outlined in Chapter 2, the species delimitation of *T. loheri* is presently unclear and this taxon might be part of a species complex in the Philippines: *T. loheri* sensu lato (s.l.; (Pelser et al., 2016b; Chapter 2). In order to be able to use *T. loheri* for conservation genetic studies in Cebu, it is important that the species delimitation of the *T. loheri* s.l. complex is resolved, because the population genetics methods that are required for these studies can only provide meaningful results if the individuals included in the datasets are conspecific.

The research presented in Chapter 2 aimed to resolve the *T. loheri* s.l. complex in the Philippines by finding groups of individuals that are morphologically distinct in leaf shape and several other vegetative characters and that therefore could be putative species under a unified species concept (De Queiroz, 2007) using morphological distinction as evidence of lineage separation. PCoA, PCA, k-means clustering and model-based clustering analyses of these characters for Philippine specimens ascribed to *T. loheri* and the morphologically similar *T. philippinense* Merr.*, T. stenophyllum* Merr.*,* and *T. trifoliolatum* Merr. revealed the absence of morphological discontinuities among these taxa. The results therefore do not support the existence of more than one species within *T. loheri* s.l. in the Philippines.

Furthermore, they showed no evidence to support that *T. philippinense, T. stenophyllum*, and Philippine specimens of *T. trifoliolatum* form separate species from *T. loheri*. However, because reproductive characters (i.e., inflorescence, flower, and fruit morphology) could not be examined in this study due to the lack of sufficient specimens with reproductive structures, it remains possible that the complex is composed of more than one species.

In this chapter, species delimitation using Generalized Mixed Yule Coalescent (GMYC; Pons et al., 2006; Monaghan et al., 2009; Fujisawa & Barraclough, 2013) and Poisson Tree Processes (PTP; Zhang et al., 2013) methods applied to nuclear and plastid DNA sequence phylogenies are used to further investigate the species delimitation of *T. loheri* s.l. GMYC and PTP are two commonly used species delimitation approaches (e.g., Papadopoulou et al., 2009; Fontaneto et al., 2011; Obertegger et al., 2012; Arribas et al., 2013; Múrria et al., 2013; Cottontail et al., 2014; Le Ru et al., 2014; Modica et al., 2014; Lang et al., 2015; Toussaint et al., 2015; Larson et al., 2016; Song et al., 2016; Blair & Bryson Jr, 2017; Mello et al., 2018; Dalstein et al., 2019) that are considered useful in exploring the presence of possible undetected species diversity within taxonomically complex groups (Carstens et al., 2013). Under a phylogenetic species concept (Baum and Shaw, 1995), the GMYC and the PTP methods identify putative species by discriminating between interspecific processes (speciation) and intraspecific processes (coalescence into a population) along the branches of a phylogenetic tree using information from branching rates. GMYC aims to find nodes in ultrametric phylogenies that mark the point of transition between speciation and coalescence, whereby nodes before this transition reflect speciation events and more recent nodes indicate coalescence events (Pons et al., 2006; Fontaneto et al., 2007; Fujisawa & Barraclough, 2013). Lineages that form putative species are marked by these points of transition. The PTP method does not require ultrametric trees. It models speciation and coalescence events directly using the branching rates (Zhang et al., 2013) and uses heuristic algorithms to classify the branches into those indicating species level processes and those indicating population level processes (Zhang et al., 2013; Tang et al., 2014). PTP assumes that the number of substitutions between species is significantly higher than the number of substitutions within species, resulting in differences between branch lengths associated with speciation events and those that are associated with coalescence events (Zhang et al., 2013).

The phylogenetic relationships among species potentially included in the *T. loheri* s.l. complex, i.e. *T. diepenhorstii* (Miq.) Latiff*, T. loheri*, *T. philippinense, T. stenophyllum,* and *T. trifoliolatum,* are still unclear. To date, *T. philippinense* and *T. stenophyllum* have not been

included in phylogenetic analyses and different studies resulted in conflicting patterns of relationships among *T. diepenhorstii* and *T. loheri*. The results of a phylogenetic study of *Tetrastigma* using four plastid DNA regions (Chen et al., 2011) suggest that *T. diepenhorstii, T. loheri*, and *T. trifoliolatum* are not each other's closest relatives. Instead, they were each resolved in different clades composed of different species (Chen et al., 2011). *Tetrastigma trifoliolatum* was not included in two subsequent phylogenetic studies using DNA sequences of a larger number of plastid regions (Wen et al., 2013; Habib et al., 2017), but both confirmed the relatively distant relationship between *T. diepenhorstii* and *T. loheri.* In contrast to the three aforementioned studies, a molecular phylogenetic study (Pelser et al., 2016b) aimed at resolving relationships among Philippine *Tetrastigma* species and using sequences of the internal transcribed spacer region (ITS) in addition to those of the same four plastid DNA regions that Chen et al. (2011) used suggested a much closer relationship between *T. diepenhorstii* and *T. loheri*. Instead of Philippine specimens identified as *T. loheri* being placed with the Indonesian *T. loheri* specimen (*Wen 10202*) that was included in the studies of Chen et al. (2011), Wen et al. (2013), and Habib et al. (2017), these specimens from the Philippines were resolved in a clade that also contains the Philippine specimen of *T. diepenhorstii* (*Wen 8261*) that was included in the phylogenies that Chen et al. (2011) and Wen et al. (2013) published. This clade is only distantly related to the two specimens of *T. trifoliolatum* that Chen et al. (2011) included in their analyses (Pelser et al., 2016b).

The study presented in this chapter aims to advance our understanding of the relationships between the putative members of the *T. loheri s.l.* complex by expanding the DNA sequence dataset of Pelser et al. (2016b). The relationships among and within the main clades within the *T. loheri* s.l. complex in this previous study were poorly resolved. The present study, therefore, also aims to improve the phylogenetic resolution of the complex to enhance the performance of the GMYC and PTP analyses (Fujisawa & Barraclough, 2013; Zhang et al., 2013). For this, DNA sequence data of the external transcribed spacer (ETS) region were added to the dataset. In addition, DNA sequences of specimens collected from the type localities of *T. loheri*, *T. philippinense, T. stenophyllum*, and *T. trifoliolatum,* and morphologically similar to these species as described in their protologues (Gagnepain, 1910; Merrill, 1912, 1914, 1916), were included to inform taxonomic conclusions. Furthermore, sequences of specimens of *T. loheri* s.l. from the four remaining forest areas of Cebu that are the focus of the research presented in Chapter 4 were added to the dataset to determine if they form a genetically distinct lineage. This was done to confirm that the specimens included in the conservation genetic research presented in this thesis are likely to be conspecific.

In this chapter, species are delineated using the unified species concept of De Queiroz (2007), which defines species as (segments of) "separately evolving metapopulation lineages" using monophyly and morphological distinction as evidence of lineage separation. Because it is not uncommon that different species delimitation investigations result in incongruent results (Carstens et al., 2013), a conservative approach to species delimitation is used in which putative GMYC and PTP-delimited species are only accepted as species if they are diagnosably morphologically distinct. Morphology data from Chapter 2 will be used for this purpose. In Chapter 2, these morphological characters were analysed using unsupervised methods to identify unknown groups in *T. loheri* s.l. that are morphologically distinct. In this chapter, a supervised method (i.e. Random Forest analysis; Breiman, 2001; Liaw & Wiener, 2002) will instead be used which will be informed by the results of the GMYC and PTP analyses. In contrast to the unsupervised analyses, this approach might be able to find subtle morphological differences between putative species that are morphologically cryptic.

3.3 OBJECTIVE

The present study aims to determine if the *T. loheri* s.l. complex in the Philippines is composed of more than one species by identifying monophyletic groups in DNA sequence phylogenies that are identified as putative species by Generalized Mixed Yule Coalescent (GMYC) and the Poisson Tree Processes (PTP) species delimitation models and that are diagnosably distinct in leaf shape and other vegetative morphological characters.

3.4 METHODOLOGY

3.4.1 Specimen sampling

DNA sequence data of 86 specimens (63 specimens of *T. loheri* s.l. and 23 specimens of other species of *Tetrastigma* and genera of Vitaceae as outgroups) were used in this study (Appendix 6). A total of 65 of these specimens were included in a previous study (Pelser et al., 2016b). To increase sampling coverage for *T. loheri* s.l. in the Philippines, 21 of the 86 specimens were collected for this study in regions that were not previously sampled. These included specimens collected from Cebu and Leyte, and specimens collected from the type localities of *T. loheri*, *T. philippinense*, *T. trifoliolatum,* and *T. stenophyllum* in Benguet, Bataan, Leyte, and La Union respectively. Type specimens of these taxa were not available for this study and specimens from their type localities that morphologically conform to these species as circumscribed in their protologues are therefore used instead. Leaf samples were

collected from each specimen and were preserved in silica gel to dehydrate the leaves and preserve their DNA. The voucher specimens were deposited in several herbaria in the Philippines and New Zealand: National Museum of the Philippines (PNH), University of the Philippines in Los Baños (CAHUP), University of the Philippines in Quezon City (PUH), University of San Carlos (CEBU), and the University of Canterbury in New Zealand (CANU).

3.4.2 **DNA extraction**

Approximately 6 mg of dried leaf tissue was placed in tubes with two metal beads and shaken for 2 minutes in an Oscillating Mill MM400 (Retsch GmbH, Haan, Germany) to pulverize the tissue. Total genomic DNA was extracted from these powdered leaf parts using a QIAGEN DNeasy Plant mini kit (Germantown, Maryland, USA) following the manufacturer's protocol.

3.4.3 **DNA regions and primers**

Two nuclear (ITS and ETS) and five plastid regions (*atpB-rbcL*, *psbA-trnH*, *rps16*, *trnL*, and *trnL-F*) were selected for DNA sequencing. The ITS and the plastid regions were PCR amplified and sequenced with the primers that were used by Pelser et al. (2016b; Appendix 7). The ETS region was sequenced using a new set of primers developed for the present study (Appendix 7).

3.4.4 **Development of ETS primers for *T. loheri*.**

ETS primers were developed from a previously generated whole genome sequence accession of *Vitis vinifera* (Vitaceae) available from GenBank (www.ncbi.nlm.nih.gov/genbank/; Appendix 8). In order to locate the ETS region within this accession, it was aligned in Geneious 6.1.8 (Biomatters, Auckland, New Zealand) with ETS GenBank sequences of ten Saxifragales species (an order closely related to Vitales; Appendix 8) and sequences of the ETS primers developed by Okuyama et al. (2004) and Baldwin and Markos (1998) (Appendix 7). Primers were subsequently designed for the segment of the *V. vinifera* accession that contains the 5'ETS region using Primer3 v 0.4.0 (Koressaar & Remm, 2007; Untergrasser et al., 2012) in Geneious using the following settings: included region 80–700 bp, target region 150–600 bp, product size 600–800 bp. This resulted in four different forward primers (84F, 85F, 86F, 87F; Appendix 7) and one reverse primer 699R. Of the four primer pairs, the 86F and 699R primer pair showed the best PCR amplification results and was initially used to amplify the ETS region of *T. loheri* s.l. specimens. However, because

PCR amplification or sequencing with this primer pair failed for some *Tetrastigma* specimens, more specific primers for *Tetrastigma* were designed using a consensus sequence of successfully generated ETS sequences of four *Tetrastigma loheri* s.l. specimens. For this, the following settings were used for Primer3 in Geneious: the included region and target region were both set at 1–615 bp (entire region) and the desired product size was confined between 600 and 615 bp. This resulted in the development of four new forward primers (3F, 4F, 5F, 6F) and one reverse primer (615R; Appendix 7). The best sequencing results were obtained using the 3F and 615R pair.

3.4.5 **PCR amplification and purification and DNA sequencing**

A total of seven DNA regions were PCR amplified using the following settings and conditions. Successful amplification was checked by running 2µl of the PCR products on a 1% agarose gel stained with SYBR Safe (Invitrogen) in 1x sodium borate buffer (Brody & Kern, 2004). The gel was subsequently observed and photographed using a Syngene G: BOXEF2 imager.

3.4.5.1 ETS

A PCR cocktail containing 1 µl of DNA, 2x Kappa plant PCR buffer containing dNTPs and MgCl2, 25 pmol/µl of forward and reverse primers (Appendix 7), and 2.5U/µl of Kappa3G Taq polymerase was prepared and brought to a volume of 15 µl with nuclease free water. The PCR conditions were set as follows: an initial denaturation temperature of 94°C for 2 minutes, denaturation temperature of 94°C for 30 seconds, annealing temperature of 51°C for 30 seconds, extension temperature of 72°C for 1 minute for 35 cycles, and a final extension of 72°C for 10 minutes.

3.4.5.2 ITS

A PCR cocktail containing 1 µl of DNA, 5x Promega Go Taq Flexi Green, 2.5 mM dNTPs, 25 pmol/µl of forward and reverse primers (Appendix 7), 25 mM of $MgCl_2$, and 0.12 µl of Go Taq polymerase was prepared and brought to a volume of 15 µl with nuclease free water. The PCR conditions were as follows: 94°C initial denaturation for 3 minutes followed by 94°C for 1 minute, 52°C for 1 minute, 72°C for 1 minute for 35 cycles, and a final extension of 72°C for 3 minutes.

3.4.5.3 psbA-trnH

A PCR cocktail containing 1 µl of DNA, 2x Kappa Ready mix, 25 pmol/µl of forward and reverse primers (Appendix 7) was prepared and brought to a volume of 10 µl with nuclease free water. The PCR conditions were as follows: initial denaturation of 94°C for 5 minutes, denaturation temperature of 94°C for 30 seconds, annealing temperature of 58°C for 30 seconds, extension temperature of 72°C for 50 seconds for 35 cycles, and a final extension of 72°C for 5 minutes.

3.4.5.4 trnL

A PCR cocktail containing 1 µl of DNA, 2x Kappa Ready Mix, 25 pmol/µl of forward and reverse primers (Appendix 7) was prepared and brought to a volume 15 µl with nuclease free water. The PCR conditions were as follows: initial denaturation of 97°C for 3 minutes, 94°C for 20 seconds, 50°C for 30 seconds, 72°C for 40 seconds for 37 cycles, and a final extension of 72°C for 5 minutes.

3.4.5.5 trnL-F

A PCR mix containing 1 µl of DNA, 2x Kappa plant PCR buffer, 25 pmol/mol of forward and reverse primers (Appendix 7), 2.5U/µl Kappa 3G Plant DNA polymerase was prepared and brought to volume of 25 µl with nuclease free water. The following PCR conditions were set: initial denaturation at 95°C for 10 minutes, followed by 20 seconds at 95°C, 15 seconds at 50°C, 30 seconds at 72°C for 40 cycles, and a final extension for 30 seconds at 72°C.

3.4.5.6 atpB-rbcL

A PCR mix containing 1 µl of DNA, 5x Promega GoTaq Flexi Green, 2.5 mM of dNTPs, 25 pmol/µl of forward and reverse primers (Appendix 7), 25mM of $MgCl_2$, 0.18 µl of Taq was prepared and brought to a volume of 15 µl with nuclease free water. PCR conditions were as follows: initial denaturation at 97°C for 3 minutes, followed by 20 seconds at 94°C, 30 seconds at 50°C, 40 seconds at 72°C for 37 cycles, a final extension at 72°C for 5 minutes.

3.4.5.7 rps16

A PCR mix containing 1 µl of DNA, 2x Kappa plant PCR buffer, 25 pmol/µl of forward and reverse primers (Appendix 7), 2.5U/µl of Kappa 3G Plant DNA polymerase was prepared and diluted to a volume of 25 µl with nuclease free water. PCR conditions were set as follows: initial denaturation at 95°C for 10 minutes, followed by 20 seconds at 95°C, 15

seconds at 50°C, 30 seconds at 72°C for 40 cycles and a final extension at 72°C for 30 seconds.

The PCR product was purified using the Promega Wizard® SV Gel and PCR Clean-Up System (Madison, Wisconsin, USA) following the manufacturer's protocol. A cycle sequencing mix containing 1 µl of purified PCR product, 25 pmol of forward and reverse primers, 0.3 µl of BigDye® Terminator v3.1 Ready Reaction Mix (ThermoFisher, New Zealand), and 2 µl of 5x BigDye® sequencing buffer (Thermo Fisher, New Zealand) was prepared and diluted to 10 µl with nuclease free water. The PCR conditions for the amplification and cycle sequencing were set as follows: initial denaturation at 96°C for 1 minute, followed by 10 seconds at 96°C, 5 seconds at 50°C, 4 minutes at 60°C for 25 cycles.

The sequencing product was mixed with 1 µl of sodium-EDTA buffer (1.5 M sodium acetate/0.25 M EDTA) and 40 µl of 95% ethanol. The mixture was centrifuged for 30 minutes at 3220 g and the supernatant was discarded afterwards. A total of 175 µl of 70% ethanol was added to the residue containing the DNA and the mixture was centrifuged again using the same speed for 5 minutes, discarding the supernatant afterwards. After it was allowed to dry for 15 minutes, the DNA was resuspended in 15 µl of Hi-Di™ Formamide (ThermoFisher, New Zealand) and was sequenced on an ABI 3130xL Genetic Analyzer at the University of the Canterbury. Geneious was used for editing the sequencing trace files.

3.4.6 **DNA sequence alignment and phylogeny reconstruction**

A total of 370 accessions (Appendix 9) were used to compile five datasets for phylogenetic analysis: ITS (83 accessions), ETS (62 accessions), the combined ETS-ITS data (nuclear dataset; 145 accessions), the combined plastid data (*atpB-rbcL*, *psbA-trnH*, *rps16*, *trnL*, and *trnL-F*; 225 accessions), and the combined nuclear and plastid datasets (370 accessions). Of these 370 accessions, 106 were newly generated for this study and 264 were obtained from GenBank. Sequences of *Ampelocissus*, *Causonis*, and *Cayratia* (Vitaceae) were used to root the *Tetrastigma* phylogenies. *Causonis* was used as an outgroup for the ITS tree. *Ampelocissus* was used to root the ETS tree. Species of *Causonis* and *Ampelocissus* were used to root the combined ETS-ITS tree. Seven species of *Ampelocissus*, *Causonis*, and *Cayratia* were used as outgroups for the plastid tree and that of the combined nuclear and plastid dataset.

DNA sequences were aligned using the Geneious Alignment method in Geneious. Insertions and deletions (indels) were subsequently coded as binary characters with Gapcode.py v.2.1 (distributed by Richard Ree, Field Museum, Chicago, Illinois, USA,

http://www.bioinformatics.org/~rick/software.html), which uses the simple indel coding method of Simmons & Ochoterena (2000).

jModelTest2 2.1.6 (Guindon & Gascuel, 2003; Darriba et al., 2012) was used on the Cyberinfrastructure for Phylogenetic Research Science Gateway v3.3 portal (CIPRES; https://www.phylo.org; Miller et al., 2010) to select a model of DNA substitution for each dataset. The number of substitution schemes to be considered was set at three. Using the Akaike Information Criterion (AIC), GTR+G was selected as the model for all datasets except *psbA-trnH,* for which the HKY+G model was used. The Markov k model (Lewis, 2001) was used for indel characters.

Bayesian Inference (BI) analysis as implemented in MrBayes 3.2.6 (Ronquist et al., 2012) was used for phylogeny reconstruction on the CIPRES cluster. The Markov Chain Monte Carlo (MCMC) analyses were set for 100,000,000 generations and were performed using four chains with a temperature setting of 0.001 and one tree saved per 100 generations. The analyses were set to stop when the average deviation of the split frequencies between both simultaneous analyses reached a value less than 0.01, suggesting convergence. The first 25% of the sampled trees were discarded as burn-in. The consensus trees resulting from the BI analyses were visualised using Fig Tree v.1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/). Clades with posterior probabilities (pp) of at least 0.95 are considered strongly supported.

3.4.7 **Ultrametric tree generation**

The sequence alignment of the combined nuclear and plastid dataset was used to generate an ultrametric tree for the GMYC analysis. This dataset was selected because it resulted in the phylogenetic tree with the highest resolution (see Results). In order to select the optimal combination of clock model and tree prior for generating this ultrametric tree, the sequence alignment was divided into two partitions in BEAUti 2.5.1 (Drummond et al., 2012): partition 1 contained the ETS, ITS, *atpB-rbcL*, *rps16*, *trnL* and *trnL-F* regions, while partition 2 only contained the *psbA-trnH* region. This was done to specify a different model of substitution for each partition following the results of the previously mentioned jModeltest analyses: the GTR model for partition 1 and the HKY substitution model for partition 2. The site model was set to Gamma with a category count of four and a substitution rate of 1.0. All monophyletic groups in the phylogenetic tree were constrained to ensure that the topology of the phylogeny remained unchanged after ultrametric conversion. This was done by supplying this topology in Newick format to the Starting Tree panel in BEAUti. The multiple monophyletic constraint was added as a prior and the same Newick tree was used to indicate

which groups of samples needed to be kept as monophyletic. Four different xml files were created in BEAUti to correspond to four different combinations of clock model and tree prior: 1] the strict clock model with the Yule model as prior (SY), 2] the Strict clock model with the coalescent constant population as prior (SC), 3] the relaxed clock log normal model with the Yule model as prior (RY), and 4] the relaxed clock log normal with the coalescent constant population as prior (RC).

A nested sampling approach (Russel et al., 2018) was employed to determine the optimal combination of clock model and tree prior. Nested sampling estimates marginal likelihood (ML) scores of the selected models and priors. The combination of model and tree prior with the highest ML score was assumed to be the optimal combination for the dataset provided. To incorporate a script for nested sampling analysis to proceed, each of the xml files that were previously created was modified using the MCMC to NS editor application found in BEAST 1.10.4 (Drummond & Rambaut, 2007; Suchard & Rambaut, 2009), using the following settings: a particle count of 10 with a subchain length of 10,000 and an Epsilon value of 1.0E-6. The modified xml files were run in BEAST on CIPRES. The highest likelihood score was obtained by RC (-15074.38) followed by SC (-15113.852). To determine if the RC combination is significantly better than the SC combination, the log Bayes factor (BF; Kass & Raftery, 1995) was obtained by computing the difference between the two log marginal likelihood scores. The minimum BF to indicate significance was computed using 2*square root of (sd1n/asd1+sd2n/asd2), where sd refers to the standard deviation estimated for each run (Bouckaert, 2019). The BF between RC and SC was 39.47 which is larger than the minimum BF of 10.74, indicating that the ML of the RC combination is significantly higher than that of the SC combination of clock model and tree prior.

The xml file with the RC combination of clock model and tree prior was run in BEAST on CIPRES to generate the ultrametric tree for the GMYC analyses. The MCMC simulation chain length was set to 50 M generations. Sampling was done at every 1000th generation. Tracer 1.7.1 (Rambaut et al., 2018) was used to inspect the effective sample sizes for each parameter. These were all at least 304, suggesting convergence. Trees were summarised using Tree Annotator v.2.5.1 with a burn-in percentage of 10% using the maximum clade credibility tree as the target tree type and common ancestor heights as the node heights.

3.4.8 **Generalized Mixed Yule Coalescent species delimitation analysis**

The ultrametric tree obtained from the combined nuclear and plastid dataset was used as input for the GMYC analysis. Outgroups were removed during the analysis because this improves the performance of the species delimitation analysis (e.g., García-Melo et al., 2019). There were three identical sequences found in the dataset, but these were not removed, because identical sequences from different individuals can improve the performance of GMYC modelling by preventing over-splitting (Talavera et al., 2013).

The GMYC method was implemented using the statistical package *splits* (Ezard et al., 2009) in R studio version 1.1.423 (RStudio Team, 2016). The function *gmyc* using a single threshold method was employed. This identifies the point that marks the transition between speciation and coalescence events (Pons et al., 2006; Fujisawa & Barraclough, 2013). The function *gmyc.support* was used to provide support values for the nodes. A likelihood ratio test, which is a built-in statistical analysis in the *gmyc* function, was used for a subset of the dataset that is composed of only *T. loheri* s.l. specimens to test if the hypothesis that there is more than one species in the *T. loheri* s.l. complex is significantly better supported than the hypothesis that this complex is only composed of a single species.

3.4.9 **Poisson Tree Processes species delimitation analysis**

The non-ultrametric phylogenetic tree generated from the combined nuclear and plastid dataset was used for the PTP analysis. It was converted into a Newick file using FigTree v.1.4.2 and was uploaded to the bPTP server (https://species.h-its.org/ptp/) which provides a Bayesian implementation of the PTP model for species delimitation (Zhang et al. 2013). The outgroups were removed from the analysis, because this has been shown to provide more accurate results (Zhang et al., 2013). The three identical sequences were retained in the dataset. The following settings were used: the number of MCMC generations was set to 500,000, thinning was set to 100, and burn-in was 10%. The likelihood scores were visually inspected for convergence which was indicated by consistent high likelihood scores in the trace file (Zhang et al., 2013). The accuracy of species delimitation at transition nodes was indicated by the value of their posterior probability (Zhang et al., 2013).

3.4.10 **Morphological support for GMYC and PTP groups**

Random Forest analyses (Breiman, 2001; Liaw & Wiener, 2002) using the combined morphology dataset (leaf shape and other vegetative morphological characters) from Chapter 2 were carried out to determine if the groups delimited by GMYC and PTP (i.e. putative species) are morphologically diagnosable using these data.

A Random Forest analysis generates multiple decision (classification) trees from bootstrap samples of the original dataset and selects the tree that best classifies the samples. A cross-validation test is performed by classifying the out-of-the bag data (OOB, data that were not sampled during bootstrapping) using the decision trees generated by the analysis. Results of the validation test are aggregated and the OOB estimate of the error rate is computed.

A dataset containing a subsample of 50 *T. loheri* s.l. specimens was compiled and used for the Random Forest analyses. This dataset includes only specimens for which morphological data was generated in Chapter 2. Specimens that were included in the phylogenetic analyses, but for which morphological data was not available, as well as single individuals that were delimited as species by GMYC and PTP, were not included in this dataset. The Random Forest analyses were run in R studio using the *randomForest* statistical package (Liaw & Wiener, 2002). The analyses were executed using ntrees=100,000 and mtry=11. The optimal mtry was determined by obtaining the square root of the total number of variables. Since the dataset has an unbalanced number of specimens in each group, the number of specimens to be drawn from each group was set to the number of specimens of the smallest group to give equal weight to each group: three for the GMYC and two for the PTP-defined groups.

After inspecting the results of the phylogenetic analyses and the GMYC and PTP analyses, two alternative species-delimitation hypotheses were explored by performing Random Forest analyses. These were aimed at testing morphological support for larger, more encompassing monophyletic groups than those identified by the GMYC and PTP analyses. Hypothesis 1 is a delimitation in which three species are recognized within the *T. loheri* s.l. complex (Groups A s.s., the *T. loheri* s.l. core clade, and H; as delimited in the Results section of this chapter). Hypothesis 2 recognizes two species within the complex (Groups H and A s.s. + the *T. loheri* s.l. core clade). Groups A s.s. and H are consistently delimited by the various GMYC and PTP analyses as putative species (Figures 3.1–3.5) and both form strongly supported clades in the combined nuclear and plastid tree (see Results). Although neither recognised by GMYC nor by PTP as a putative species, the *T. loheri* s.l. core clade is a large monophyletic group that is strongly supported and is composed of all *T. loheri* s.l. specimens that are not included in Groups A s.s. and H (see Results). Specimens of Group A s.s. are consistently resolved as more closely related to the *T. loheri* s.l. core clade than to Group H (Figures 3.1–3.5). Three specimens were drawn for Random Forest analysis for Hypothesis 1, and four for Hypothesis 2.

3.5 RESULTS

3.5.1 Phylogenetic analyses

The ITS was the most informative region, with 910 characters of which 52 were gap-coded and 329 (36%) showed unique site patterns. This was followed by the following datasets in decreasing order of unique site patterns: the nuclear dataset (combined ETS-ITS) with 1587 characters of which 118 were gap-coded and 510 (32%) showed unique site patterns, ETS with 658 characters of which 48 were gap-coded and 181(28%) showed unique site patterns, the combined nuclear and plastid dataset had 5594 characters of which 444 were gap-coded and 1494 (27%) showed unique site patterns, and, finally, the concatenated plastid dataset with 3983 of which 303 were gap coded and 985(25%) showed unique site patterns.

The Bayesian consensus trees of all datasets resulted in similar topologies with respect to the relationships among the *T. loheri* s.l. complex and the other species that were included in the analyses, except for some differences in the relationships among *T. ellipticum* Merr.*, T. scariosum* Planch.*, T. glabratum* Planch.*, T. magnum* Merr., and *T. harmandii* Planch. (Figures 3.1–3.5).

Nine groups of *T. loheri* s.l. specimens were composed of accessions that either formed clades in phylogenies obtained from all or most datasets or that were at least consistently placed in each other's phylogenetic vicinity: A, B, C, D, E, F, G, H, and I (Figures 3.1–3.5). These nine groups together formed a strongly supported (i.e. pp >0.95; ETS; Figure 3.2) or poorly supported (i.e. pp <0.95; plastid dataset; Figure 3.4) clade, formed a polytomy with clades composed of specimens from other taxa than *T. loheri* s.l. (combined ETS-ITS and combined nuclear and plastid dataset; Figures 3.3 and 3.5), or formed a poorly supported non-monophyletic *T. loheri* s.l. group (ITS; Figure 3.1).

Tetrastigma loheri s.l. group H was monophyletic with strong support in all trees (Figures 3.1–3.5). It was comprised of individuals from the islands of Leyte, Mindanao, and Panay. In all trees, it was resolved as being more distantly related to the other *T. loheri* s.l. groups than these are to each other. This pattern was strongly supported.

Group A was strongly supported as monophyletic by all datasets (Figures 3.1–3.3, 3.5) except by the plastid dataset (Figure 3.4). Group A was composed of three individuals from Mindanao (Group A s.s.). A specimen from Indonesia (*Chen & Lu 158*: T_diepenhorstii_ChenLu158) sequenced by Habib et al. (2017) was nested within Group A in the combined nuclear and plastid tree (Figure 3.5) and formed a grade with the other Group A

specimens in the plastid tree (Figure 3.4). ITS and ETS sequences of this specimen were not available and it was therefore not included in the nuclear datasets.

Groups B, C, D, E, F, G and I formed a clade in consensus trees obtained from all datasets. This '*T. loheri* s.l. core clade' was strongly supported in all trees, except for the plastid tree (Figure 3.4). Group B was monophyletic with strong support in the ITS and combined ETS-ITS tree (Figures 3.1 and 3.3) but with weak support in the combined nuclear and plastid tree (Figure 3.5). Its specimens were placed in a polytomy with members of other groups in the ETS and plastid trees (Figures 3.2 and 3.4). Group B consisted of individuals from Cagayan and Ilocos Norte, both of which were regions in the north of Luzon.

Group C was monophyletic with weak support in the combined ETS-ITS tree and strong support in the combined nuclear and plastid tree (Figures 3.3 and 3.5). Its specimens were placed in a polytomy with a specimen of *T. philippinense* from the type locality of this species in Benguet (central Luzon) in the ITS and ETS tree, forming a strongly (ITS) or weakly (ETS) supported clade (Figures 3.1 and 3.2). *Tetrastigma philippinense* formed a weakly supported clade with Group C specimens in the combined ETS-ITS tree (Figure 3.3). The Group C specimens were placed in a polytomy with a large number of members of other *T. loheri* s.l. groups in the plastid tree (Figure 3.4). Group C specimens were mostly from the central region of Luzon (Nueva Ecija).

Group E was a strongly supported clade in the ITS and the combined ETS-ITS trees (Figures 3.1 and 3.3). Its specimens formed a grade in the ETS tree (Figure 3.2) and were part of a large polytomy in the plastid phylogeny (Figure 3.4). Group E plants were from the eastern region of Luzon and this group contained a Philippine specimen that was identified as *T. diepenhorstii* by the collector (*Wen 8261:* T_diepenhorstii8261; Figures 3.4 and 3.5). Only plastid sequences were generated for this specimen by Chen et al. (2011) and it was therefore not included in the nuclear tree. In the combined nuclear and plastid tree, the other Group E specimens formed a weakly supported clade which was placed in a polytomy with *Wen 8261* and other specimens (Figure 3.5).

Two of the three specimens of Group G formed a strongly supported clade in all trees, except for the plastid tree (Figure 3.4). These two specimens were both collected at the type locality of *T. trifoliolatum* on the island of Leyte. The third specimen, from the nearby island of Samar, only formed a weakly supported clade with the two other specimens in the combined nuclear and plastid tree (Figures 3.5). In addition to the two specimens from the type locality of *T. trifoliolatum*, plastid sequences of two specimens from Peninsular Malaysia (*Wen 8350:* T_trifoliolatum8350) and from New Guinea (*Wen 10758:* T_

trifoliolatum10758) generated by Chen et al. (2011) and identified as *T. trifoliolatum* were included in the plastid and combined nuclear and plastid datasets (Figures 3.4 and 3.5). These two specimens were resolved as only distantly related to the two Philippine specimens ascribed to this species and were instead placed in a strongly supported clade with specimens of *T. crenatum* Jackes and *T. papillosum* Planch.

Groups D, F, and I were not resolved as monophyletic (Figures 3.1–3.5). Group D contained specimens from the central to western regions of Luzon and included specimens that were collected from the type localities of *T. loheri*, *T. philippinense*, and *T. stenophyllum*. Group F plants were mostly from the island of Cebu. Group I plants were collected from the islands of Negros and Panay.

Of the five datasets that were analysed, the combined nuclear and plastid dataset resulted in the most well resolved phylogeny (Figure 3.5). This dataset was therefore used for the species-delimitation analyses.

3.5.2 **GMYC and PTP species delimitation results**

The number of species identified in the GMYC analysis was 33 with a confidence interval of 11–41. Within the *T. loheri* s.l. complex, GMYC identified 12 species. Nine of these were composed of more than one specimen (Figure 3.5). Groups B, C, E, G, H, and I were all delimited as species by the GMYC modelling. Three additional putative species groups identified in the analysis each closely corresponded to Groups A, D, and F (here referred to as Groups A s.s., D s.s., and F s.s.), but did not include one of their specimens, which were individually segregated as species. Of the other *Tetrastigma* species that were included in the combined nuclear and plastid dataset, only *T. harmandii, T. papillosum*, and *T. scariosum* were recovered in the GMYC analyses. The delimited species all had AIC-based support values of $p < 0.95$. A likelihood ratio test that was conducted with a dataset that only contains *T. loheri* s.l. specimens to specifically test the hypothesis that this complex is composed of more than one species resulted in a likelihood value of 597.1079 for the hypothesis that there are no distinct species groups within *T. loheri* s.l. and a value of 601.3039 for the hypothesis that more than one species group exists. The latter hypothesis was significantly better supported ($p=0.015$).

The results of the PTP delimitation estimated a mean number of species of 32.55 with a confidence interval between 17 and 61. There were nine species delimited by PTP within *T. loheri* s.l. which consisted of six groups and three individual specimens (Figure 3.5). The six groups were BCD, FI, A s.s., E, H, and G s.s. Group BCD was the combination of Groups B,

C, and D s.s. Likewise, Group FI was the combination of F and I. Of the remaining species included in the dataset, only *T. harmandii* and *T. lawsonii* (King) Burkill were identified by PTP. None of the PTP-delimited species was strongly supported (i.e. pp >=0.95).

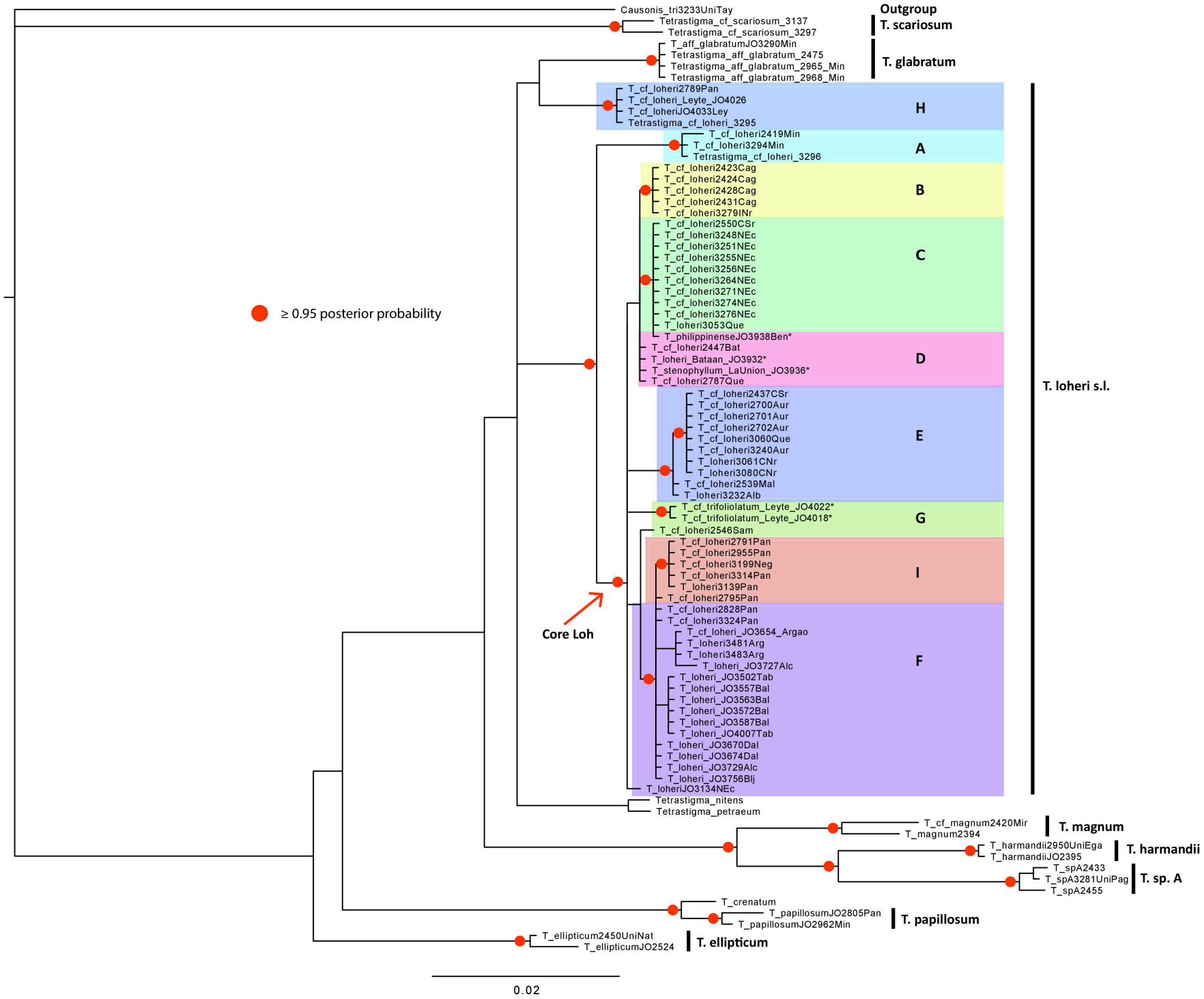

Causonis_tri3233UniTay
Tetrastigma_cf_scariosum_3137
Tetrastigma_cf_scariosum_3297
Outgroup
T. scariosum
T_aff_glabratumJO3290Min
Tetrastigma_aff_glabratum_2475
Tetrastigma_aff_glabratum_2965_Min
Tetrastigma_aff_glabratum_2968_Min
T. glabratum
T_cf_loheri2789Pan
T_cf_loheri_Leyte_JO4026
T_cf_loheriJO4033Ley
Tetrastigma_cf_loheri_3295
H
T_cf_loheri2419Min
T_cf_loheri3294Min
Tetrastigma_cf_loheri_3296
A
T_cf_loheri2423Cag
T_cf_loheri2424Cag
T_cf_loheri2428Cag
T_cf_loheri2431Cag
T_cf_loheri3279INr
B
T_cf_loheri2550CSr
T_cf_loheri3248NEc
T_cf_loheri3251NEc
T_cf_loheri3255NEc
T_cf_loheri3256NEc
T_cf_loheri3264NEc
T_cf_loheri3271NEc
T_cf_loheri3274NEc
T_cf_loheri3276NEc
T_loheri3053Que
C
≥ 0.95 posterior probability
T_philippinenseJO3938Ben*
T_cf_loheri2447Bat
T_loheri_Bataan_JO3932*
T_stenophyllum_LaUnion_JO3936*
T_cf_loheri2787Que
D
T. loheri s.l.
T_cf_loheri2437CSr
T_cf_loheri2700Aur
T_cf_loheri2701Aur
T_cf_loheri2702Aur
T_cf_loheri3060Que
T_cf_loheri3240Aur
T_loheri3061CNr
T_loheri3080CNr
T_cf_loheri2539Mal
T_loheri3232Alb
E
T_cf_trifoliolatum_Leyte_JO4022*
T_cf_trifoliolatum_Leyte_JO4018*
G
T_cf_loheri2546Sam
T_cf_loheri2791Pan
T_cf_loheri2955Pan
T_cf_loheri3199Neg
T_cf_loheri3314Pan
T_loheri3139Pan
I
T_cf_loheri2795Pan
Core Loh
T_cf_loheri2828Pan
T_cf_loheri3324Pan
T_cf_loheri_JO3654_Argao
T_loheri3481Arg
T_loheri3483Arg
T_loheri_JO3727Alc
T_loheri_JO3502Tab
T_loheri_JO3557Bal
T_loheri_JO3563Bal
T_loheri_JO3572Bal
T_loheri_JO3587Bal
T_loheri_JO4007Tab
T_loheri_JO3670Dal
T_loheri_JO3674Dal
T_loheri_JO3729Alc
T_loheri_JO3756Blj
F
T_loheriJO3134NEc
Tetrastigma_nitens
Tetrastigma_petraeum
T_cf_magnum2420Mir
T_magnum2394
T. magnum
T_harmandii2950UniEga
T_harmandiiJO2395
T. harmandii
T_spA2433
T_spA3281UniPag
T_spA2455
T. sp. A
T_crenatum
T_papillosumJO2805Pan
T_papillosumJO2962Min
T. papillosum
T_ellipticum2450UniNat
T_ellipticumJO2524
T. ellipticum
0.02

*Figure 3.1. The internal transcribed spacer (ITS) Bayesian inference tree. Clades with posterior probabilities of at least 0.95 are marked with red circles. Accessions marked with * indicate specimens collected from type localities. Highlighted groups are discussed in the text. 'Core Loh' indicates the* T. loheri *s.l. core clade.*

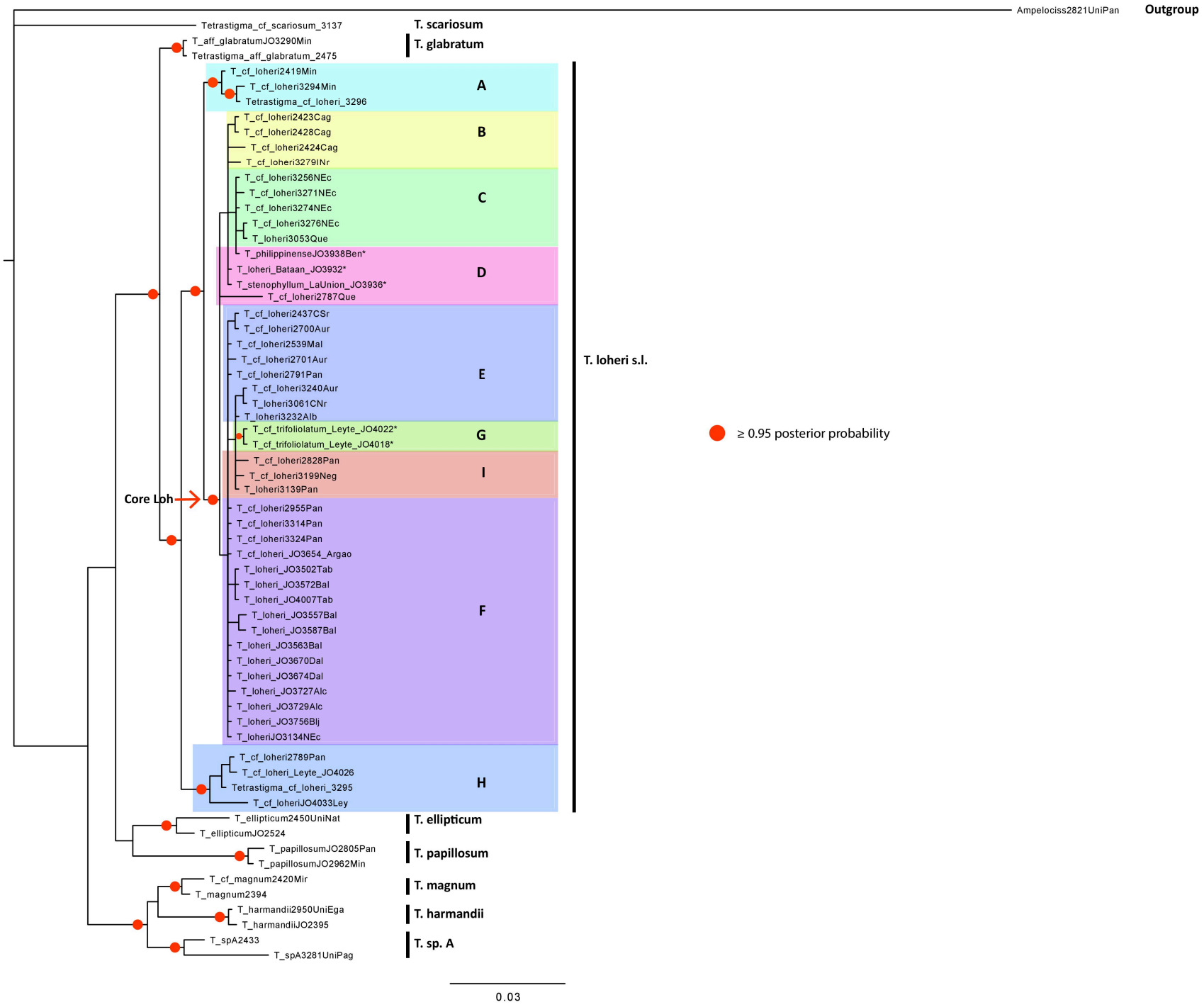

Ampelociss2821UniPan
Outgroup
Tetrastigma_cf_scariosum_3137
T. scariosum
T_aff_glabratumJO3290Min
Tetrastigma_aff_glabratum_2475
T. glabratum
T_cf_loheri2419Min
T_cf_loheri3294Min
Tetrastigma_cf_loheri_3296
A
T_cf_loheri2423Cag
T_cf_loheri2428Cag
T_cf_loheri2424Cag
T_cf_loheri3279INr
B
T_cf_loheri3256NEc
T_cf_loheri3271NEc
T_cf_loheri3274NEc
T_cf_loheri3276NEc
T_loheri3053Que
C
T_philippinenseJO3938Ben*
T_loheri_Bataan_JO3932*
T_stenophyllum_LaUnion_JO3936*
T_cf_loheri2787Que
D
T_cf_loheri2437CSr
T_cf_loheri2700Aur
T_cf_loheri2539Mal
T_cf_loheri2701Aur
T_cf_loheri2791Pan
T_cf_loheri3240Aur
T_loheri3061CNr
T_loheri3232Alb
E
T_cf_trifoliolatum_Leyte_JO4022*
T_cf_trifoliolatum_Leyte_JO4018*
G
T_cf_loheri2828Pan
T_cf_loheri3199Neg
T_loheri3139Pan
I
Core Loh
T_cf_loheri2955Pan
T_cf_loheri3314Pan
T_cf_loheri3324Pan
T_cf_loheri_JO3654_Argao
T_loheri_JO3502Tab
T_loheri_JO3572Bal
T_loheri_JO4007Tab
T_loheri_JO3557Bal
T_loheri_JO3587Bal
T_loheri_JO3563Bal
T_loheri_JO3670Dal
T_loheri_JO3674Dal
T_loheri_JO3727Alc
T_loheri_JO3729Alc
T_loheri_JO3756Blj
T_loheriJO3134NEc
F
T_cf_loheri2789Pan
T_cf_loheri_Leyte_JO4026
Tetrastigma_cf_loheri_3295
T_cf_loheriJO4033Ley
H
T. loheri s.l.
≥ 0.95 posterior probability
T_ellipticum2450UniNat
T_ellipticumJO2524
T. ellipticum
T_papillosumJO2805Pan
T_papillosumJO2962Min
T. papillosum
T_cf_magnum2420Mir
T_magnum2394
T. magnum
T_harmandii2950UniEga
T_harmandiiJO2395
T. harmandii
T_spA2433
T_spA3281UniPag
T. sp. A
0.03

*Figure 3.2. The external transcribed spacer (ETS) Bayesian inference tree. Clades with posterior probabilities of at least 0.95 are marked with red circles. Accessions marked with * indicate specimens collected from type localities. Highlighted groups are discussed in the text. 'Core Loh' indicates the* T. loheri *s.l. core clade.*

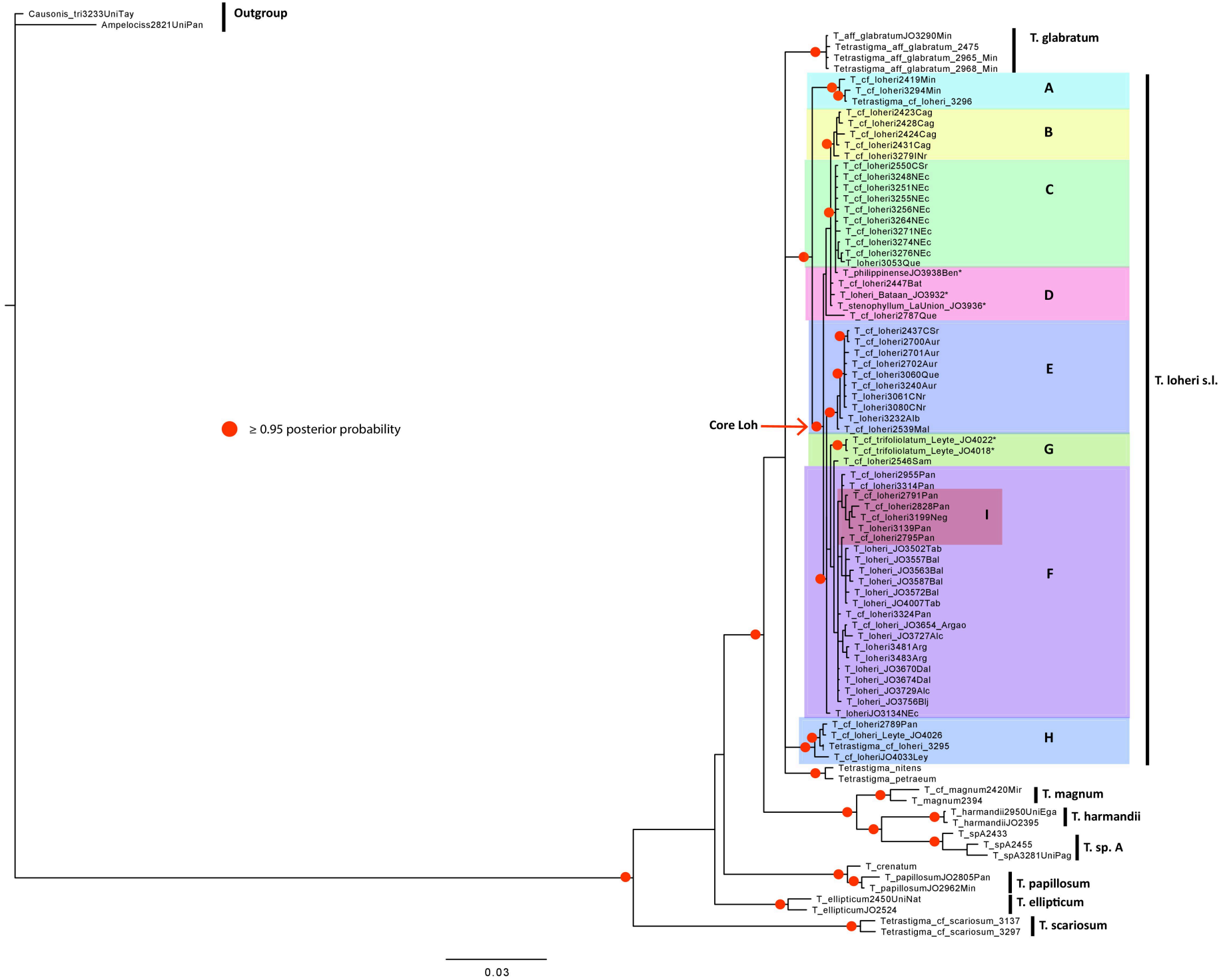

Causonis_tri3233UniTay
Ampelociss2821UniPan
Outgroup
≥ 0.95 posterior probability
Core Loh
T_aff_glabratumJO3290Min
Tetrastigma_aff_glabratum_2475
Tetrastigma_aff_glabratum_2965_Min
Tetrastigma_aff_glabratum_2968_Min
T. glabratum
T_cf_loheri2419Min
T_cf_loheri3294Min
Tetrastigma_cf_loheri_3296
A
T_cf_loheri2423Cag
T_cf_loheri2428Cag
T_cf_loheri2424Cag
T_cf_loheri2431Cag
T_cf_loheri3279INr
B
T_cf_loheri2550CSr
T_cf_loheri3248NEc
T_cf_loheri3251NEc
T_cf_loheri3255NEc
T_cf_loheri3256NEc
T_cf_loheri3264NEc
T_cf_loheri3271NEc
T_cf_loheri3274NEc
T_cf_loheri3276NEc
T_loheri3053Que
C
T_philippinenseJO3938Ben*
T_cf_loheri2447Bat
T_loheri_Bataan_JO3932*
T_stenophyllum_LaUnion_JO3936*
T_cf_loheri2787Que
D
T_cf_loheri2437CSr
T_cf_loheri2700Aur
T_cf_loheri2701Aur
T_cf_loheri2702Aur
T_cf_loheri3060Que
T_cf_loheri3240Aur
T_loheri3061CNr
T_loheri3080CNr
T_loheri3232Alb
T_cf_loheri2539Mal
E
T_cf_trifoliolatum_Leyte_JO4022*
T_cf_trifoliolatum_Leyte_JO4018*
G
T_cf_loheri2546Sam
T_cf_loheri2955Pan
T_cf_loheri3314Pan
T_cf_loheri2791Pan
T_cf_loheri2828Pan
T_cf_loheri3199Neg
I
T_loheri3139Pan
T_cf_loheri2795Pan
T_loheri_JO3502Tab
T_loheri_JO3557Bal
T_loheri_JO3563Bal
T_loheri_JO3587Bal
T_loheri_JO3572Bal
T_loheri_JO4007Tab
T_cf_loheri3324Pan
T_cf_loheri_JO3654_Argao
T_loheri_JO3727Alc
T_loheri3481Arg
T_loheri3483Arg
T_loheri_JO3670Dal
T_loheri_JO3674Dal
T_loheri_JO3729Alc
T_loheri_JO3756Blj
F
T_loheriJO3134NEc
T_cf_loheri2789Pan
T_cf_loheri_Leyte_JO4026
Tetrastigma_cf_loheri_3295
T_cf_loheriJO4033Ley
H
T. loheri s.l.
Tetrastigma_nitens
Tetrastigma_petraeum
T_cf_magnum2420Mir
T_magnum2394
T. magnum
T_harmandii2950UniEga
T_harmandiiJO2395
T. harmandii
T_spA2433
T_spA2455
T_spA3281UniPag
T. sp. A
T_crenatum
T_papillosumJO2805Pan
T_papillosumJO2962Min
T. papillosum
T_ellipticum2450UniNat
T_ellipticumJO2524
T. ellipticum
Tetrastigma_cf_scariosum_3137
Tetrastigma_cf_scariosum_3297
T. scariosum
0.03

*Figure 3.3. The combined external transcribed spacer (ETS) and internal transcribed spacer (ITS) Bayesian inference tree. Clades with posterior probabilities of at least 0.95 are marked with red circles. Accessions marked with * indicate specimens collected from type localities. Highlighted groups are discussed in the text. 'Core Loh' indicates the* T. loheri *s.l. core clade.*

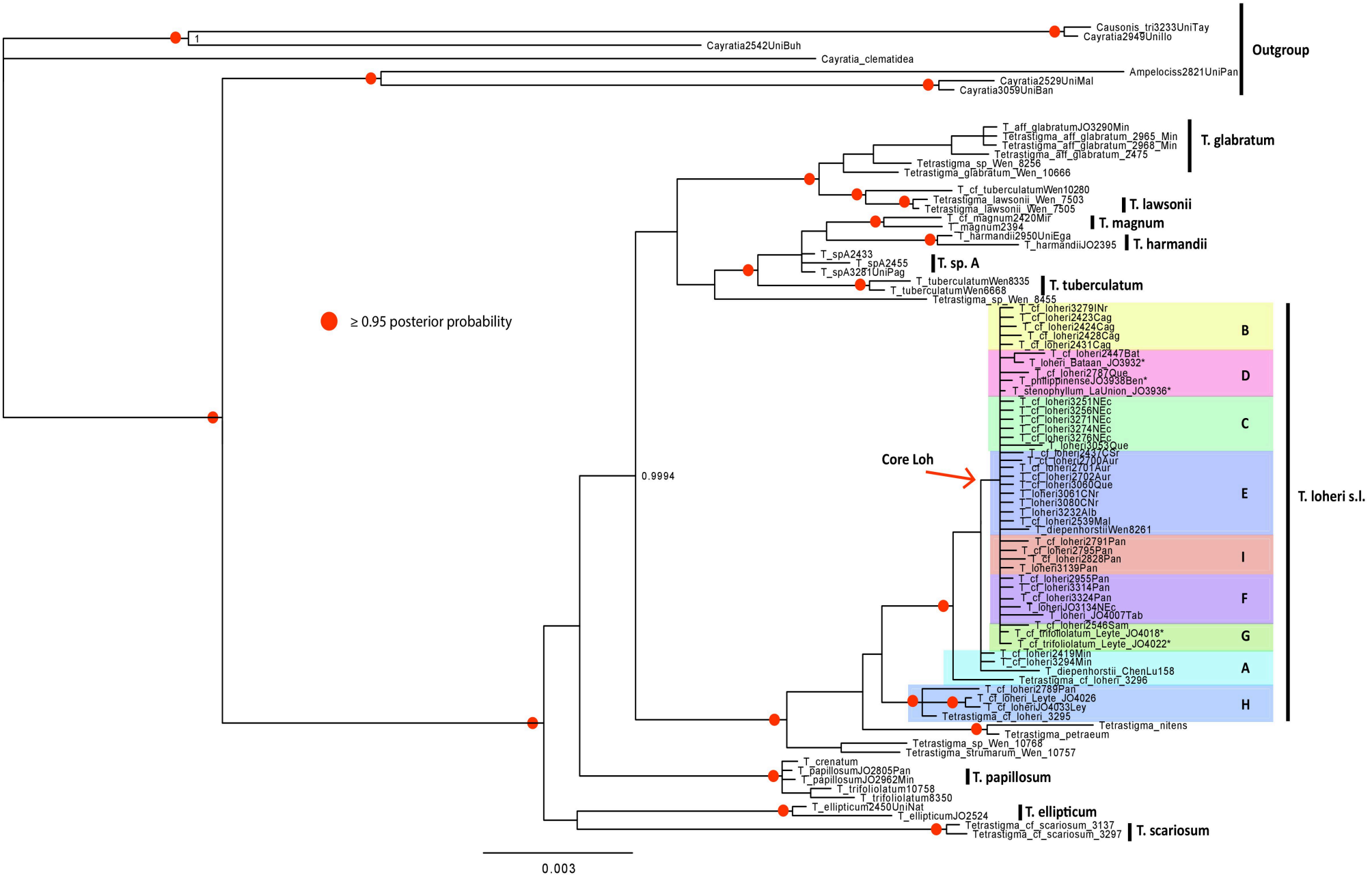

Causonis_tri3233UniTay
Cayratia2949Unillo
Outgroup
1
Cayratia2542UniBuh
Cayratia_clematidea
Ampelociss2821UniPan
Cayratia2529UniMal
Cayratia3059UniBan
T_aff_glabratumJO3290Min
Tetrastigma_aff_glabratum_2965_Min
Tetrastigma_aff_glabratum_2968_Min
Tetrastigma_aff_glabratum_2475
T. glabratum
Tetrastigma_sp_Wen_8256
Tetrastigma_glabratum_Wen_10666
T_cf_tuberculatumWen10280
Tetrastigma_lawsonii_Wen_7503
Tetrastigma_lawsonii_Wen_7505
T. lawsonii
T_cf_magnum2420Mir
T. magnum
T_magnum2394
T_harmandii2950UniEga
T_harmandiiJO2395
T. harmandii
T_spA2433
T_spA2455
T_spA3281UniPag
T. sp. A
T_tuberculatumWen8335
T_tuberculatumWen6668
T. tuberculatum
Tetrastigma_sp_Wen_8455
≥ 0.95 posterior probability
T_cf_loheri3279INr
T_cf_loheri2423Cag
T_cf_loheri2424Cag
T_cf_loheri2428Cag
T_cf_loheri2431Cag
B
T_cf_loheri2447Bat
T_loheri_Bataan_JO3932*
T_cf_loheri2787Que
T_philippinenseJO3938Ben*
T_stenophyllum_LaUnion_JO3936*
D
T_cf_loheri3251NEc
T_cf_loheri3256NEc
T_cf_loheri3271NEc
T_cf_loheri3274NEc
T_cf_loheri3276NEc
T_loheri3053Que
C
Core Loh
T_cf_loheri2437CSr
T_cf_loheri2700Aur
T_cf_loheri2701Aur
T_cf_loheri2702Aur
T_cf_loheri3060Que
T_loheri3061CNr
T_loheri3080CNr
T_loheri3232Alb
T_cf_loheri2539Mal
T_diepenhorstiiWen8261
E
T. loheri s.l.
0.9994
T_cf_loheri2791Pan
T_cf_loheri2795Pan
T_cf_loheri2828Pan
T_loheri3139Pan
I
T_cf_loheri2955Pan
T_cf_loheri3314Pan
T_cf_loheri3324Pan
T_loheriJO3134NEc
T_loheri_JO4007Tab
F
T_cf_loheri2546Sam
T_cf_trifoliolatum_Leyte_JO4018*
T_cf_trifoliolatum_Leyte_JO4022*
G
T_cf_loheri2419Min
T_cf_loheri3294Min
T_diepenhorstii_ChenLu158
Tetrastigma_cf_loheri_3296
A
T_cf_loheri2789Pan
T_cf_loheri_Leyte_JO4026
T_cf_loheriJO4033Ley
Tetrastigma_cf_loheri_3295
H
Tetrastigma_nitens
Tetrastigma_petraeum
Tetrastigma_sp_Wen_10768
Tetrastigma_strumarum_Wen_10757
T_crenatum
T_papillosumJO2805Pan
T_papillosumJO2962Min
T. papillosum
T_trifoliolatum10758
T_trifoliolatum8350
T_ellipticum2450UniNat
T_ellipticumJO2524
T. ellipticum
Tetrastigma_cf_scariosum_3137
Tetrastigma_cf_scariosum_3297
T. scariosum
0.003

*Figure 3.4. The plastid tree inferred by Bayesian inference using combined data from atpB-rbcL , psbA-trnH , rps16, trnL and trnL-F. Clades with posterior probabilities of at least 0.95 are marked with red circles. Accessions marked with * indicate specimens collected from type localities. Highlighted groups are discussed in the text. 'Core Loh' indicates the* T. loheri *s.l. core clade.*

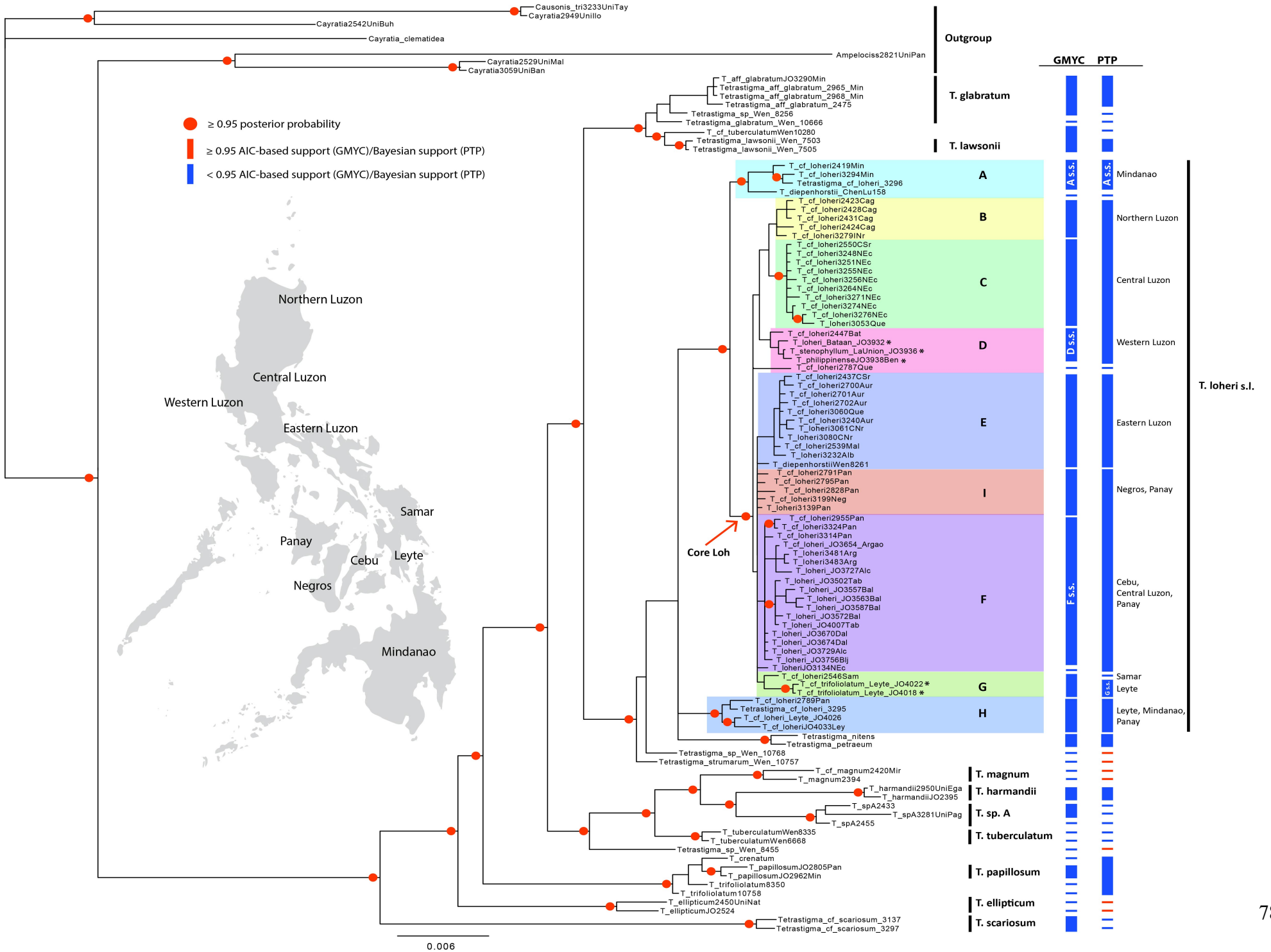

Causonis_tri3233UniTay
Cayratia2949Unillo
Cayratia2542UniBuh
Cayratia_clematidea
Ampelociss2821UniPan
Cayratia2529UniMal
Cayratia3059UniBan
Outgroup
GMYC
PTP
≥ 0.95 posterior probability
≥ 0.95 AIC-based support (GMYC)/Bayesian support (PTP)
< 0.95 AIC-based support (GMYC)/Bayesian support (PTP)
T. glabratum
T_aff_glabratumJO3290Min
Tetrastigma_aff_glabratum_2965_Min
Tetrastigma_aff_glabratum_2968_Min
Tetrastigma_aff_glabratum_2475
Tetrastigma_sp_Wen_8256
Tetrastigma_glabratum_Wen_10666
T. lawsonii
T_cf_tuberculatumWen10280
Tetrastigma_lawsonii_Wen_7503
Tetrastigma_lawsonii_Wen_7505
A
T_cf_loheri2419Min
T_cf_loheri3294Min
Tetrastigma_cf_loheri_3296
T_diepenhorstii_ChenLu158
A s.s.
Mindanao
B
T_cf_loheri2423Cag
T_cf_loheri2428Cag
T_cf_loheri2431Cag
T_cf_loheri2424Cag
T_cf_loheri3279INr
Northern Luzon
C
T_cf_loheri2550CSr
T_cf_loheri3248NEc
T_cf_loheri3251NEc
T_cf_loheri3255NEc
T_cf_loheri3256NEc
T_cf_loheri3264NEc
T_cf_loheri3271NEc
T_cf_loheri3274NEc
T_cf_loheri3276NEc
T_loheri3053Que
Central Luzon
D
T_cf_loheri2447Bat
T_loheri_Bataan_JO3932 *
T_stenophyllum_LaUnion_JO3936 *
T_philippinenseJO3938Ben *
T_cf_loheri2787Que
D s.s.
Western Luzon
T. loheri s.l.
E
T_cf_loheri2437CSr
T_cf_loheri2700Aur
T_cf_loheri2701Aur
T_cf_loheri2702Aur
T_cf_loheri3060Que
T_cf_loheri3240Aur
T_loheri3061CNr
T_loheri3080CNr
T_cf_loheri2539Mal
T_loheri3232Alb
T_diepenhorstiiWen8261
Eastern Luzon
I
T_cf_loheri2791Pan
T_cf_loheri2795Pan
T_cf_loheri2828Pan
T_cf_loheri3199Neg
T_loheri3139Pan
Negros, Panay
Core Loh
F
T_cf_loheri2955Pan
T_cf_loheri3324Pan
T_cf_loheri3314Pan
T_cf_loheri_JO3654_Argao
T_loheri3481Arg
T_loheri3483Arg
T_loheri_JO3727Alc
T_loheri_JO3502Tab
T_loheri_JO3557Bal
T_loheri_JO3563Bal
T_loheri_JO3587Bal
T_loheri_JO3572Bal
T_loheri_JO4007Tab
T_loheri_JO3670Dal
T_loheri_JO3674Dal
T_loheri_JO3729Alc
T_loheri_JO3756Blj
T_loheriJO3134NEc
F s.s.
Cebu,
Central Luzon,
Panay
G
T_cf_loheri2546Sam
T_cf_trifoliolatum_Leyte_JO4022 *
T_cf_trifoliolatum_Leyte_JO4018 *
G s.s.
Samar
Leyte
H
T_cf_loheri2789Pan
Tetrastigma_cf_loheri_3295
T_cf_loheri_Leyte_JO4026
T_cf_loheriJO4033Ley
Leyte, Mindanao,
Panay
Tetrastigma_nitens
Tetrastigma_petraeum
Tetrastigma_sp_Wen_10768
Tetrastigma_strumarum_Wen_10757
T. magnum
T_cf_magnum2420Mir
T_magnum2394
T. harmandii
T_harmandii2950UniEga
T_harmandiiJO2395
T. sp. A
T_spA2433
T_spA3281UniPag
T_spA2455
T. tuberculatum
T_tuberculatumWen8335
T_tuberculatumWen6668
Tetrastigma_sp_Wen_8455
T. papillosum
T_crenatum
T_papillosumJO2805Pan
T_papillosumJO2962Min
T_trifoliolatum8350
T_trifoliolatum10758
T. ellipticum
T_ellipticum2450UniNat
T_ellipticumJO2524
T. scariosum
Tetrastigma_cf_scariosum_3137
Tetrastigma_cf_scariosum_3297
0.006
Northern Luzon
Central Luzon
Western Luzon
Eastern Luzon
Samar
Panay
Cebu
Leyte
Negros
Mindanao

*Figure 3.5. Phylogenetic tree inferred by Bayesian inference using the concatenated nuclear and plastid dataset. Clades with posterior probabilities of at least 0.95 are marked with red circles. Accessions marked with * indicate plants collected from type localities. Highlighted groups are discussed in the text. 'Core Loh' indicates the* T. loheri *s.l. core clade. Species delimited by the Generalized Mixed Yule Coalescent (GMYC) and the Poisson Tree Processes (PTP) methods are indicated by blue (<0.95 AIC-based support/posterior probability) and red bars (=>0.95 AIC-based support/posterior probability). Narrow delimitations (sensu stricto) of Groups A, D, F, and G are designated as A s.s., D s.s., F s.s., and G s.s. respectively in the corresponding bars. The islands or the regions on an island where the GMYC or PTP groups are found are indicated and are shown in the map.*

3.5.3 **Random Forest analyses**

Random Forest analyses were used to determine if the putative species delimited with the GMYC and PTP methods are diagnosably distinct when leaf shape and other vegetative morphological data of Chapter 2 are used. These analyses resulted in a 64% out-of-the-bag (OOB) error of classifying samples to GMYC groups within the *T. loheri*i s.l. complex. The confusion matrix (Table 3.1) shows that between 17–100% of the specimens of each GMYC group was misclassified by the model, except for specimens of Group H (0% classification error). The OOB estimate of error of classifying samples to *T. loheri* s.l. groups as delimited by PTP using the morphological data was 68%. Except for Group H (0% percent classification error), each PTP group had a classification error of between 60–100% (Table 3.2).

Alternative species-delimitation Hypothesis 1 (Groups A s.s., the *T. loheri* s.l. core clade, and H) had a 26 % OOB error rate and Hypothesis 2 (Groups H and A s.s. + the *T. loheri* s.l. core clade) had a 10% OOB error rate.

Table 3.1. Confusion matrix for putative Tetrastigma loheri *s.l. species delimited by the Generalized Mixed Yule Coalescent method. The rows refer to the actual groups. The columns refer to the predicted groups. Class.error refers to percentage classification error.*

Groups	A s.s.	B	C	D s.s.	E	F s.s.	G	H	I	class.error
A s.s.	0	0	0	0	0	0	1	1	1	100%
B	0	2	0	0	2	1	0	0	0	60%
C	1	1	4	1	0	0	0	0	3	60%
D s.s.	0	0	1	0	1	0	1	0	1	100%
E	0	2	0	1	3	0	0	3	1	70%
F s.s.	0	0	0	0	0	5	1	0	0	17%
G	0	1	0	0	0	1	0	0	1	100%
H	0	0	0	0	0	0	0	4	0	0%
I	0	0	2	1	0	0	2	0	0	100%

Table 3.2. Confusion matrix for putative Tetrastigma loheri *s.l. species delimited by the Poisson Tree Process method. The rows refer to the actual groups. The columns refer to the predicted groups. Class.error refers to percentage classification error.*

Groups	A s.s.	BCD	E	FI	G s.s.	H	class.error
A s.s.	0	1	0	0	0	2	100%
BCD	1	6	5	5	2	0	68%
E	0	3	4	0	1	2	60%
FI	1	5	0	2	4	0	83%
G s.s.	1	0	1	0	0	0	100%
H	0	0	0	0	0	4	0%

Table 3.3. Confusion matrix for putative Tetrastigma loheri *s.l. species based on alternative Hypothesis 1 (see text). The rows refer to the actual groups. The columns refer to the predicted groups. Class.error refers to percentage classification error.*

	core Loh	A s.s.	H	class.error
core Loh	34	7	2	21%
A s.s.	1	0	2	100%
H	1	0	3	25%

Table 3.4. Confusion matrix for putative Tetrastigma loheri *s.l. species based on alternative Hypothesis 2 (see text). The rows refer to the actual groups. The columns refer to the predicted groups. Class.error refers to percentage classification error.*

	core Loh + A.s.s.	H	class.error
core Loh + A.s.s.	42	4	9%
H	1	3	25%

3.6 DISCUSSION

This study aimed to determine if the *T. loheri* s.l. complex in the Philippines is composed of more than one species by identifying morphologically distinct monophyletic groups in DNA sequence phylogenies that are identified as putative species by Generalized Mixed Yule Coalescent and the Poisson Tree Processes species delimitation models.

3.6.1 Phylogenetic analyses

The trees reconstructed from the five datasets showed similar phylogenetic patterns. There were no issues when the five datasets were combined for phylogenetic analyses because topological incongruences among the trees were not well supported (e.g., relationships among *T. magnum, T. harmandii,* and *T.* sp A.; Figures 3.1–3.5). The five datasets were concatenated to use all available information for resolving clades within *T. loheri* s.l. complex and between *Tetrastigma* species and this resulted in increased resolution and nodal support (Fig. 3.5).

The ITS tree (Figure 3.1) was more resolved than the ETS (Figure 3.2) and plastid trees (Figure 3.4). However, the resolution within the clades of the ITS tree was poor. With the aim of enhancing the performance of the GMYC and PTP delimitation models on the phylogenies by improving the resolution of the ITS clades (Fujisawa & Barraclough, 2013; Zhang et al., 2013), the ETS regions were therefore sequenced and added to the phylogenetic analysis which resulted in the improvement of the resolution within the clades of the ITS in the combined ETS and ITS tree (e.g., A and B; Figure 3.3.) The plastid regions were sequenced and included in the analysis to resolve relationships at the higher level of classification than species. However, only a selection of specimens was sequenced for the plastid regions because preliminary analyses with smaller datasets showed that the plastid sequence variation among specimens within these clades was too low to provide more resolution. This resulted in several missing data in the dataset for the plastid regions (Appendix 9). Despite low sequence variation of the plastid regions, a combined phylogenetic analysis of the ITS, ETS, and the plastid regions resolved clades which were never found in the separate and combined trees of the ITS and ETS (e.g., D s.s. and G; Figures 3.1–3.3, 3.5).

3.6.2 **GMYC and PTP species delimitation**

The results of a likelihood ratio test as part of a GMYC analysis of a combined nuclear and plastid DNA sequence phylogeny in which only specimens of *T. loheri* s.l. are included suggests that this complex is composed of more than one species (null hypothesis of no distinct species groups within *T. loheri* s.l. rejected: p= 0.015). In fact, the GMYC analysis of the complete dataset that also included representatives of other *Tetrastigma* species suggests that there may be 12 species within *T. loheri* s.l. (Figure 3.5). Also the results of the PTP analysis indicate the presence of more than one species within the complex (Figure 3.5). The delimitation of the nine *T. loheri* s.l. species that resulted from the PTP analysis is largely congruent with those delimited by the GMYC analysis, but the former are mostly more broadly delineated and therefore fewer in number. Whereas the GMYC method identified specimen groups B, C, and D s.s. as separate species, the PTP model grouped these together into Group BCD. Similarly, the GMYC model identified Groups I, F s.s., and a single accession of a specimen from Nueva Ecija (*Barcelona 3868 with Pelser*: T_loheriJO3134NEc) as distinct, but the results of the PTP analysis instead suggest that these specimens are conspecific (i.e. Group FI). Only Group G is more narrowly delimited by the PTP method than by the GMYC method (Figure 3.5).

Although both species delimitation methods suggest that the *T. loheri* s.l. group is composed of multiple species and similar delimitation hypotheses were obtained, the differences between these delimitations indicate that care should be taken when interpreting these results. They might suggest that the performance of the GMYC and PTP modelling was not optimal in the present study. This is further indicated by the large confidence intervals of the GMYC and PTP species estimations (i.e. GMYC 11–41 species; PTP 17–61 species). In addition, statistical support for the putative species within the *T. loheri* s.l. complex was poor in both delimitations (i.e. all AIC support and posterior probabilities < 0.95). Furthermore, only a few species of the other *Tetrastigma* species included in the analyses were correctly identified by the GMYC (i.e. *T. harmandii, T. papillosum, T. scariosum*) and PTP methods (i.e. *T. harmandii, T. lawsonii*).

3.6.3 **Biogeographic patterns**

Although the GMYC and PTP analyses did not result in a single and strongly supported species delimitation hypothesis for the *T. loheri* s.l. complex in the Philippines, some of the inferred groups might be biologically meaningful, because they are composed of specimens that are mostly from the same island (e.g., Groups A s.s., F, and I; Figure 3.5) or

from the same region within an island (e.g., Groups B, C, D; Figure 3.5). Because these islands and regions are separated by geographic features that may prevent or significantly reduce gene flow among them (e.g., tall mountain ranges or bodies of water), it is possible that the correlation between the species-delimitation groups and biogeographic patterns in the data is a result of their reproductive isolation. This could therefore mean that these groups are indeed species or incipient species.

Alternatively, the biogeographic patterns in the phylogenies might not be an indication of the existence of multiple species or incipient species within the *T. loheri* s.l. complex, but rather provide an explanation for obtaining species delimitation hypotheses that are too narrowly delineated as a result of over splitting. That is because geographic structuring of intraspecific genetic variation might result in identifying a partially genetically isolated population as a separate species (Papadopoulou et al., 2008; Fujisawa & Barraclough, 2013; Luo et al., 2018). Although reduced geneflow between some of the GMYC and PTP groups as a result of geographic barriers is an intuitively appealing explanation, especially in an archipelago like the Philippines, geographic structuring can also be explained by spatial autocorrelation: individuals of a species that are living closer together are expected to be more genetically similar than individuals that live further apart (Meirmans, 2012).

Group H is biogeographically unique among the GMYC and PTP groups, because it is composed of specimens from islands that are relatively distant from each other: Leyte, Mindanao, Panay (Figure 3.5). In addition, although Group H individuals have been found to be sympatric with *T. loheri* s.l. specimens from other groups on each of these islands, they are genetically distinct. For example, *Obico 986* (T_cf_loheri_Leyte_JO4026), which is a member of Group H, was collected in the same forest on Leyte where two of the Group G individuals (*Obico 983:* T_cf_trifoliolatum_Leyte_JO4018 and *Obico 984*: T_cf_trifoliolatum_Leyte_JO4022) were found. Although such patterns of sympatry could be a result of recent long-distance dispersal, they might also indicate reproductive isolation in sympatry and, as such, be considered as evidence for recognizing Group H as a distinct species.

3.6.4 **Finding morphological support for GMYC and PTP groups**

Following the recommendations of Talavera et al. (2013), Zhang et al. (2013), and others, GMYC and PTP-delimited groups are here considered as putative species that should be validated with other lines of evidence. This is an integrative taxonomic approach in which different lines of evidence are used to support species recognition (Sukumaran & Knowles,

2017; Luo et al., 2018). This approach is particularly important for GMYC and PTP-delimited species, because there are indications that these methods can result in false positives as a result of over splitting (Luo et al., 2018). The present study, therefore, also aimed to determine if the putative species delimited by GMYC and PTP are diagnosably distinct in leaf shape or other vegetative morphological characters. Such groups could be recognized as species under the unified species concept (De Queiroz, 2007) assuming monophyly as well as morphological distinction as evidence of lineage separation.

To determine if the GMYC and PTP-delimited groups are morphologically diagnosable, Random Forest analyses were carried out using a previously generated dataset of leaf shape and several other vegetative characters (Chapter 2). Because the Random Forest method requires groups to be composed of at least two specimens, putative species to which only one specimen was assigned (i.e. singletons) could not be included in these analyses. The four singletons identified within *T. loheri* s.l. (*Barcelona 3667 with Sarmiento*: T_cf_loheri2787Que; *Barcelona 3726 et al.*: T_cf_loheri2546Sam; *Barcelona 3868 with Pelser*: T_loheriJO3134NEc; *Chen & Lu 158*: T_diepenhorstii_ChenLu158) might represent species that are rare and additional representatives of these lineages need to be studied to clarify their taxonomic status. However, although *Chen & Lu 158* was not available for study, the results of the morphometric analyses presented in Chapter 2 do not indicate that the other three specimens are morphologically distinct from the other *T. loheri* s.l. specimens.

The overall Random Forest classification errors for the GMYC and PTP delimitation hypotheses (excluding singletons) of the *T. loheri* s.l. complex were high (64% and 68% respectively). However, this error rate decreased when a three-species hypothesis was explored (26%; Table 3.3). The two-species hypothesis had the lowest error rate (10%; Table 3.4). The specimens of Group H (Tables 3.1–3.2) obtained a 0% classification error in both GMYC and PTP confusion matrices. Although these results may suggest that group H shows some level of morphological distinction, some specimens from other groups were incorrectly classified to Group H (Tables 3.1–3.2) and a relatively high misclassification rate was still obtained when alternative species delimitations were applied in which Group H is also recognized as a putative species (Tables 3.3–3.4). Therefore, in conclusion, the vegetative characters used in this study do not appear to indicate that the putative species identified by the GMYC or PTP methods are diagnosably distinct (Tables 3.1–3.2), even when broader species delimitation hypotheses were considered (Tables 3.3–3.4).

3.6.5 **Species delimitation in the *T. loheri* s.l. complex**

The results of this study do not provide support for recognizing more than one species within the *T. loheri* s.l. complex in the Philippines under the species-delimitation criteria used, because they do not indicate the presence of monophyletic groups that are diagnosably morphologically distinct. This supports the conclusion of Chapter 2 of this thesis that this lineage is most likely a species that is very variable in its vegetative morphology and in particular in leaf shape. As pointed out in Chapter 2, this is well illustrated by the two Philippine specimens that were collected at the type locality of *T. trifoliolatum* (*Obico 983:* T_cf_trifoliolatum_Leyte_JO4018 and *Obico 984*: T_cf_trifoliolatum_Leyte_JO4022). Both were collected in each other's vicinity and are resolved as each other's closest relatives (Figures 3.1–3.3, 3.5), but they are substantially different in their vegetative morphology as indicated by their different positions in the various morphometric ordination plots (e.g., Chapter 2: Figure 2.14).

Although the available evidence suggests that the *T. loheri* s.l. might only be composed of a single species in the Philippines, it is possible that future studies of reproductive or other characters that were not included in this study will provide evidence in support of recognizing more than one species. This should be taken into account when taxonomic conclusions are drawn from this study (see below). The failure to find clear support for the existence of multiple species within this complex might also potentially be attributed to non-optimal performance of the GMYC and PTP modelling as a result of errors in phylogenetic reconstruction or a lack of phylogenetic resolution due to lack of informative characters. However, although indeed only a few clades obtained high posterior probabilities in the combined plastid and nuclear phylogeny (e.g., Groups A, H, *T. loheri* s.l. core clade; Figure 3.5), none of these was identified as a potential species by the GMYC or PTP analyses with high statistical support.

Of all possible putative species considered, most support exists for recognizing Group H as taxonomically distinct at the species level. Resolved as more distantly related to the other *T. loheri* s.l. groups than they are to each other, group H is phylogenetically distinct and forms a well-supported clade (Figures 3.1–3.5). Furthermore, it contains specimens that are sympatric with members of other groups and is therefore potentially reproductively isolated from them. In addition, it is the group for which the lowest Random Forest classification errors were obtained, at least in the GMYC and PTP confusion matrices (Tables 3.1–3.2), suggesting that there might be some morphological features in which it is different from the remainder of the *T. loheri* s.l. complex.

3.6.6 **Taxonomic implications**

Under the assumption that the *T. loheri* s.l. complex only consists of one species and considering that specimens identified as *T. diepenhorstii*, *T. loheri*, *T. philippinense, T. stenophyllum,* and *T. trifoliolatum* are resolved among its members, taxonomic changes to the species-level classification of *Tetrastigma* might be required. Such changes require certainty regarding the correct application of taxonomic names. This is problematic because type specimens were not included in this study. In an attempt to mitigate this, specimens from the type localities of four of the five species that have been reported from the Philippines and that were identified as putative members of the *T. loheri* s.l. complex because of morphological similarities (Chapter 2) were included in the analyses presented here and in Chapter 2. These specimens show the diagnostic morphological features as outlined in the species protologues (Gagnepain, 1910; Merrill, 1912, 1914, 1916). Unfortunately, specimens from the type locality of *T. diepenhorstii* in Sumatra (Miquel 1861) were unavailable for study. Specimens identified as *T. trifoliolatum* were included in this study. Two Philippine specimens (*Obico 983*: T_cf_trifoliolatum_Leyte_JO4018 and *Obico 984*: T_cf_trifoliolatum_Leyte_JO4022) are nested within Group G of the *T. loheri* s.l. complex, whereas two specimens from Peninsular Malaysia (*Wen 8350*: T_trifoliolatum8350) and New Guinea (*Wen 10758:* T_trifoliolatum10758) form a clade with specimens of *T. papillosum* and *T. crenatum* (Figure 3.5). This raises doubt about the correct identification of these specimens. Because the Philippine specimens were collected from the type locality of *T. trifoliolatum* in Leyte, it is most likely that they represent this species. Unfortunately, specimens of *Wen 8350* and *10758* at US were not accessible for this study. Hence, it was not possible to verify their identification. If *Obico 983* and *984* indeed represent *T. trifoliolatum* and if future studies would provide evidence in support of recognizing Group G as a distinct species, then the name *T. trifoliolatum* is available for this species.

Tetrastigma trifoliolatum was synonymized with *T. diepenhorstii* by Latiff (2001), who compared the types of both names and considered them to be conspecific. The two specimens identified as the latter species that were included in the analyses are indeed resolved as members of the *T. loheri* s.l. complex, although they are placed in different groups (i.e. Groups A and E; Figure 3.5). One specimen (*Chen & Lu 158*: T_diepenhorstii_ChenLu158), which was identified as a singleton, is from Indonesia (Bogor Botanical Garden; Appendix 10) and the other (*Wen 8261:* T_diepenhorstii8261) was collected in the Philippines. *Wen 8261* could not be located at US, but the vegetative morphology *Chen & Lu 158* is in agreement with the description in the protologue of *T. diepenhorstii* (Miquel, 1861). It is

therefore possible that the type of *T. diepenhorstii* is indeed a member of the *T. loheri* s.l. complex, although this remains to be confirmed.

Specimens from the type localities of *T. loheri*, *T. philippinense*, and *T. stenophyllum* are resolved as closely related to each other. They are placed in Group D s.s. (GMYC) or Group BCD (PTP; Figure 3.5). Under the assumption that these specimens indeed represent these three species, *T. loheri*, *T. philippinense*, and *T. stenophyllum* are therefore synonyms even if the most narrow species-delimitation hypothesis provided by the GMYC and PTP analyses is adopted.

If *T. diepenhorstii*, *T. loheri*, *T. philippinense, T. stenophyllum,* and *T. trifoliolatum* are indeed conspecific, *T. diepenhorstii* is the correct name for *T. loheri* s.l., because this name has priority (Turland et al., 2018). However, because neither the type specimen nor specimens from the type locality of *T. diepenhorstii* were included and because this name has not been used for Philippine plants, other than for *Wen 8261* (Pelser et al., 2011 onwards), which was not available for study, I will use the name *T. loheri* to refer to the Philippine representatives of this taxon in Chapter 4.

3.6.7 Implications for the use of *T. loheri* specimens for conservation genetic studies in Cebu

The present study does not provide conclusive evidence in support of recognizing more than one species within the *T. loheri* s.l. complex in the Philippines. However, even if future studies would show that this hypothesis is incorrect and that multiple species should be recognized instead, the results of the analyses presented here suggest that the 14 *T. loheri* s.l. specimens collected from Cebu are conspecific, because they are all resolved as members of the same GMYC and PTP-delimited species (Group F or IF; Figure 3.5). These 14 specimens were collected from all four forest areas that were selected for the conservation genetic study presented in Chapter 4. Therefore, I conclude that it is unlikely that the dataset compiled for this study is composed of specimens belonging to more than one species and is therefore appropriate for studying patterns of genetic diversity and connectivity among these four remaining forested areas in Cebu.

CHAPTER 4: No evidence of low genetic diversity despite high levels of inbreeding and poor genetic connectivity among *Tetrastigma loheri* populations in remaining forest areas in Cebu, Philippines

4.1 ABSTRACT

Little is known about the effects of habitat fragmentation on the patterns of genetic diversity and genetic connectivity of species in the remaining forests of Cebu, a Philippine island that has a long history of deforestation and has lost nearly all of its forest cover. To contribute towards filling this knowledge gap, I conducted a conservation genetic study. Using data from 13 microsatellite loci that I developed for *Tetrastigma loheri*, a common vine in Philippine forests, I studied patterns of genetic diversity and genetic connectivity for the four largest of the remaining forest areas in Cebu. Evidence of relatively high levels of inbreeding was found in the four areas, despite no evidence of low genetic diversity. The four areas were genetically differentiated, suggesting low genetic connectivity. Isolation by distance was found and partly explains the pattern of genetic differentiation among the four areas. The evidence of inbreeding and low genetic connectivity in a commonly encountered species such as *T. loheri* in Cebu suggests that the impact of habitat fragmentation is likely greater on rare plant species with restricted distributions in Cebu forests. Conservation recommendations for the remaining forest areas in Cebu resulting from this study include the establishment of ecological corridors between nearby areas to improve the movement of pollinators and seed dispersers between them.

4.2 INTRODUCTION

The Philippines is an archipelagic country of around 7,000 islands that support immense biological diversity (Myers et al., 2000). Around 90% of its area was covered with forest when the Spanish colonisers arrived in the mid-16th century (Westoby, 1989; Lasco et al., 2001). Over the centuries that followed, there was an accelerating loss of forest cover due to large-scale logging, agricultural expansion, and upland migration (Chokkalingam et al., 2006; Laurance, 2007). According to the Forest Management Bureau of the Philippines, as of 2015, the forest cover in the Philippines was estimated at 23% (Forest Management Bureau, 2019).

The most extensive deforestation in the Philippines was perceived by many conservation practitioners to have taken place in Cebu (Gonzalez et al., 1999; Jakosalem et

al., 2013; Paguntalan et al., 2015), one of the largest islands in the central Philippines (Figure 4.1). This was believed to have commenced during the establishment of the first Spanish colony on the island in the mid-16th century (Paguntalan et al., 2015). By 1875, only 6.6 – 11% forest cover was remaining in Cebu (Bankoff, 2007) and, according to the Forest Management Bureau (2019), the forest cover of the island of Cebu was estimated at 1.57 % in 2015. The long history of deforestation in Cebu has also resulted in severely fragmented forest (Gonzalez et al. 1999, Paguntalan et al. 2015). Remnant forest fragments are mostly confined to non-arable and steep areas of the island (Seidenschwarz, 1988; Paguntalan et al., 2015). Land conversion and wood harvesting for domestic consumption were the major contributors to forest decline on the island (Paguntalan et al., 2015). The intensive deforestation on Cebu resulted in the extinction of many of its endemic fauna (Rabor, 1959; Gonzalez et al., 1999; Paguntalan et al., 2015). For instance, four of the 12 bird subspecies endemic to Cebu are now believed to be extinct (Paguntalan & Jakosalem, 2008a) and also the Visayan warty pig (*Sus cebrifrons*) and Visayan spotted deer (*Rusa alfredi*) have been extirpated from this island (Paguntalan et al., 2015).

Despite the destruction, degradation and fragmentation of forests, biodiversity that is endemic to the Philippines is still present in the remaining forested areas of Cebu. This includes fish, butterflies, damselflies, skinks, plants, birds, and bats, and most of these species are island endemics (Gonzalez et al., 1999; Paguntalan et al., 2015). Cebu is now recognized as a centre of bird endemism in the Philippines and is considered as an important endemic bird area by Bird Life International (Paguntalan et al., 2015). The conservation of birds and bats has received considerable interest in the recent past and this resulted in the rediscovery of species that were previously thought to be extinct (Dutson et al., 1993; Magsalay, 1993; Gadiana, 2004). Much less is known about the conservation status of endemic species in other groups of organisms, such as insects and plants. Examples of the latter are Cebu cinnamon (*Cinnamomum cebuense* Kosterm.) (Cadiz & Buot Jr, 2009), *Cynometra cebuensis* F.Seid. (Seidenschwarz, 2013), *Lepeostegeres cebuensis* Barcelona, Nickrent & Pelser (Pelser et al., 2016c) and *Vaccinium cebuense* Salares & Pelser (Salares et al., 2018).

In recent years, there have been a growing number of initiatives that address the need to conserve Cebu's biodiversity in the remaining forest fragments (e.g., Paguntalan and Jakosalem (2008a); Paguntalan and Jakosalem (2008b); Cadiz and Buot Jr (2009); Jakosalem et al. (2013); Malaki et al. (2013); Lillo et al. (2015); Malaki (2016)). A large proportion of these studies have focused on species diversity studies. While these conservation studies are

valuable, published data on the genetic diversity of forest species in Cebu are, to my knowledge, not available.

Genetic diversity is one the facets of biodiversity. Genetic diversity is the variety of alleles and genotypes in populations and is characterised by parameters such as heterozygosity and allele frequency (Frankham et al., 2010). Genetic diversity is an important aspect of biodiversity because it is the source of alleles that enables populations to be resilient, i.e. to adapt to changes in their environment, thus enabling their long-term persistence (Jamieson et al., 2008; Frankham et al., 2010). Genetic diversity is associated with genetic connectivity. Genetic connectivity reflects the degree to which gametes (e.g. pollen in plants) or seeds disperse between populations (Slatkin, 1987; Lowe & Allendorf, 2010). Genetic connectivity maintains the level of genetic diversity between populations, thus preventing reduction of genetic diversity through genetic drift and inbreeding (Ellstrand & Elam, 1993; Jenkins & Stevens, 2018).

One possible reason for not explicitly addressing genetic diversity in conservation studies is the assumption that species diversity can serve as surrogate for genetic diversity since species diversity and genetic diversity are both influenced by the same environmental factors, such as habitat size and connectivity, and variation in the environment (Kahilainen et al., 2014). Although some studies provide support for a correlation between species and genetic diversity (e.g. Struebig et al. (2011); Blum et al. (2012)), other research demonstrates that this correlation does not always exist (Taberlet et al., 2012; Kahilainen et al., 2014). Because species diversity is not always a suitable proxy for genetic diversity, genetic diversity should not be neglected in biodiversity studies, but should be an integrated element. This is a prudent approach because conservation plans that have been designed to protect an area on the sole basis of its species diversity might either underestimate or overestimate the genetic diversity of the area.

Genetic diversity and genetic connectivity are negatively affected by habitat fragmentation. Fragmentation of habitats causes contraction of continuous habitat into smaller remnants resulting in decreased population sizes and increased separation between populations rendering them isolated (Young et al., 1996; Lienert, 2004). Small, isolated populations are prone to genetic diversity loss through genetic drift (Ellstrand & Elam, 1993). This is because a population of few individuals would have a high chance of losing some alleles, especially the rare ones, by chance (Ellstrand & Elam, 1993). Loss of alleles can lead to a decrease in heterozygotes and the eventual fixation of remnant alleles within the population (Ellstrand & Elam, 1993). Genetic diversity loss through drift, consequently,

increases extinction risk of the population because of the reduced number of alleles that would potentially allow the population to adapt under a changing environment. In addition, in a small isolated population, inbreeding occurs where offspring are produced as a result of mating between related individuals (Ellstrand & Elam, 1993). Inbreeding can result in inbreeding depression through an increase in homozygotes (Ellstrand & Elam, 1993). This may expose deleterious recessive alleles which can reduce reproductive fitness (Ellstrand & Elam, 1993; Lowe et al., 2005).

In the absence of published population genetic studies of forest species in Cebu, little is known about the patterns of genetic diversity and genetic connectivity of the remaining forests of this island. This information is important in conservation management that aims to maintain genetic diversity for the long-term persistence of forest species. As a first step towards filling this knowledge gap, this chapter aims to provide the first information about patterns of genetic diversity and connectivity of plants among Cebu forest areas by utilizing genetic data from one of Cebu's forest species, *Tetrastigma loheri* Gagnep. *Tetrastigma loheri* is a woody vine from the grape family (Vitaceae) and it is commonly encountered in the remaining forest areas of Cebu. It is a forest vine species, so it is dependent on forest trees, on which it climbs to reach the canopy and obtain sunlight. The reduction of forest cover resulting in smaller forest fragments is therefore expected to have had an impact on the patterns of genetic diversity and connectivity of *T. loheri*.

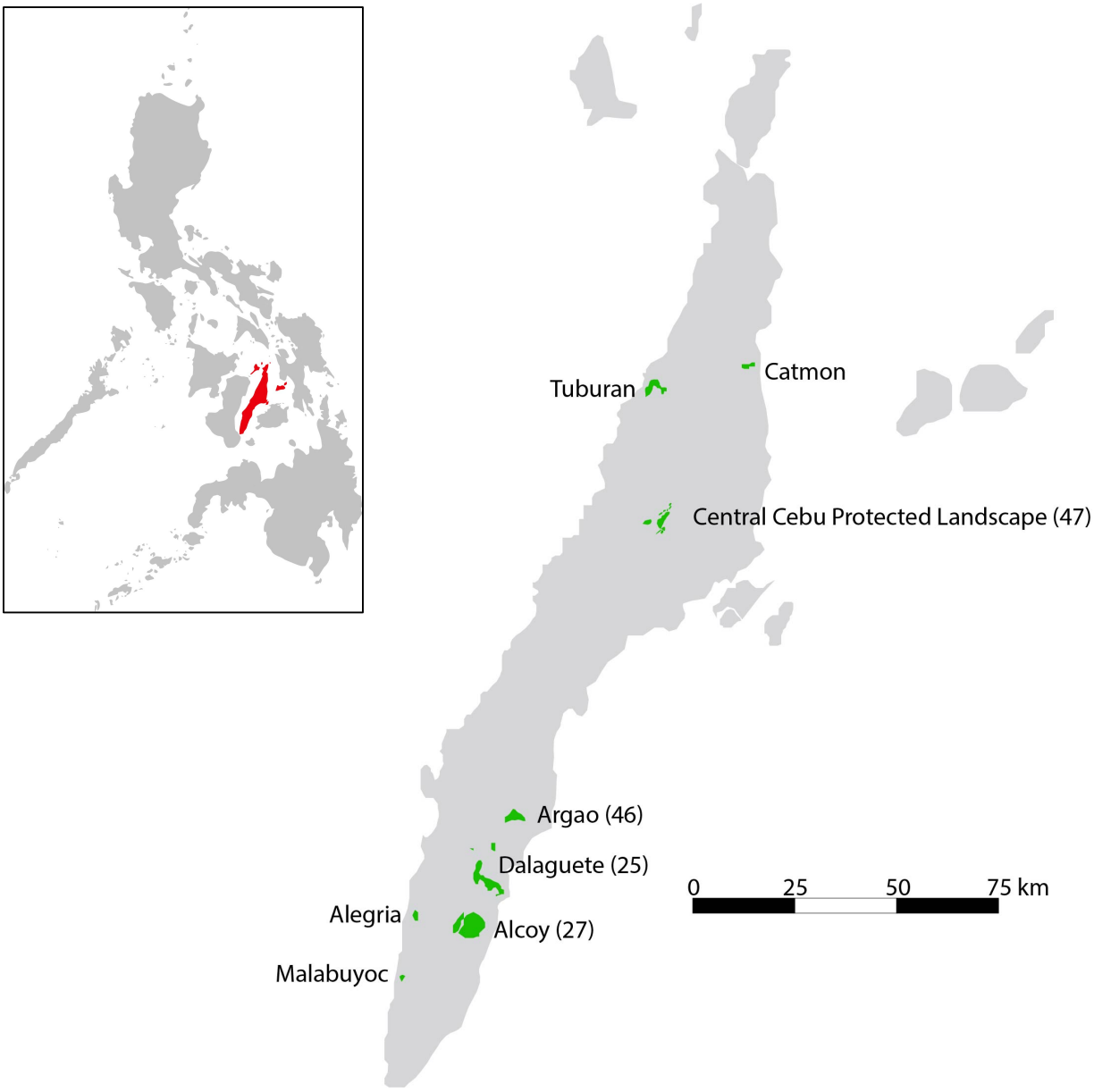


Figure 4.1. A map of Cebu island showing the location of the eight largest remaining forested areas (redrawn from Paguntalan et al. (2015) with permission from the authors). The four sampling areas for Tetrastigma loheri *are indicated by the number of collected individuals in each area. An inset shows the map of the Philippines highlighting the location of the island of Cebu in red.*

4.3 OBJECTIVES

This chapter aims to contribute to a better understanding of the patterns of genetic diversity and connectivity of four of the eight largest remaining forested areas of Cebu: Alcoy, Argao, Dalaguete, and the Central Cebu Protected Landscape (CCPL). It explores these patterns using microsatellite genotyping data for *T. loheri*, a plant species that is commonly encountered in all four areas. It specifically addresses the following questions:

a. What is the genetic diversity of *T. loheri* in Alcoy, Argao, Dalaguete, and CCPL?

b. Is inbreeding evident in Alcoy, Argao, Dalaguete, and CCPL?

c. What is the pattern of genetic connectivity among the populations of *T. loheri* in Alcoy, Argao, Dalaguete, and CCPL?

4.4 METHODOLOGY

4.4.1 Study areas

Four of the eight largest forested areas in Cebu were selected for this study: those in the municipalities of Argao, Dalaguete, and Alcoy (extending into Boljoon) in the south-eastern part of the island, and the Central Cebu Protected Landscape (CCPL) in the central region (Figure 4.1). The distances among the forests of Argao, Dalaguete, and Alcoy are approximately 10 km and the distance between these three areas and the CCPL is approximately 60–90 km.

Almost no primary forest vegetation is left in Cebu (Rabor, 1959; Seidenschwarz, 1988; Kummer et al., 1994; Gonzalez et al., 1999; Paguntalan et al., 2015). CCPL is composed of variously-sized forest fragments that together form 420 ha of low-midland forest in the municipality of Balamban and barangay Tabunan in Cebu City. Some tracts of forest in this area are estimated to be of primary growth (Gonzalez et al., 1999, Paguntalan et al. 2015), but most parts consist of secondary vegetation. The forests in the other three areas are secondary growth (Gonzalez et al., 1999; Paguntalan et al., 2015). In all four forested areas, the secondary forest is in various stages of succession. People collect firewood and clear parts of these forests for agriculture. Argao forest covers at least 507.17 ha and includes the dry secondary forest of Mt. Lantoy (Paguntalan et al., 2015) and the surrounding foothill forest, which contains stands of exotic mahogany (*Swietenia macrophylla* King) which was used for reforestation. Dalaguete has 654.3 ha of lowland to mid-montane secondary forest and also contains mahogany stands (Gonzalez et al., 1999; Paguntalan et al., 2015). Alcoy is

home to Nug-as forest, which is the largest remaining forest of Cebu with an area of 1,036 ha of mid-montane secondary vegetation (Gonzalez et al., 1999; Paguntalan et al., 2015).

Only CCPL is protected and managed by the national government under the National Integrated Protected Areas System of the Philippines. Although it does not enjoy formal protection, Nug-as forest in Alcoy is managed by a Peoples Organisation of farmers called Kapunungan sa mga Mag-uuma sa Yutang Lasangnon sa Bulolacao under a Community-based Forest Management Agreement with the Department of Environment and Natural Resources (DENR) of the Philippine government. We are not aware of any people's organisation that is involved in a similar community-based forest management program or a clear investment in conservation by local government in Argao or Dalaguete.

4.4.2 **Focal species**

Tetrastigma loheri is a species that is commonly encountered in forest ecosystems in the Philippines (Pelser et al., 2016b). It is a generalist in terms of forest habitat preference in the four areas, because it can be found in the interior as well as along the fringes of the forest, and even in small areas of regenerating tree vegetation along roadsides. *Tetrastigma loheri* is dioecious and has unisexual flowers that are white to cream. They are relatively small and form a compound cymose inflorescence (Appendix 11A). Little is known about the pollination biology of *Tetrastigma,* but it is most likely pollinated by insects (Brizicky, 1965). Honey bees are among the potential pollinators as they have been observed visiting flowers of *T. diepenhorstii* (Zakaria et al., 2017). The fruits of *T. loheri* are red, fleshy globose or elongated berries (Appendix 11B) that turn dark purple to black (Pelser et al., 2016b) suggesting bird-dispersal (Janson, 1983).

Tetrastigma loheri is morphologically variable throughout its range in the Philippines. Conclusive evidence of genetically distinct subgroups, which could be an indication that this taxon needs to be more narrowly defined, was, however, not discovered (Chapter 3). Morphological and molecular phylogenetic patterns presented in Chapter 3 further indicate that *T. loheri* plants from the four study areas in Cebu would belong to the same taxon even if future studies would show that this species needs to be more narrowly delimited, because representative samples (n= 14) were resolved in the same clade in the DNA sequence phylogeny (Figure 3.5 in Chapter 3) and they do not display large morphological variation.

Identifying *Tetrastigma* species is a challenge because of the absence of comprehensive taxonomic descriptions for most species and a taxonomic revision of the genus. Particularly the diagnostic reproductive characters of the species are poorly known

(Pelser et al., 2016b). This might be due to the dioecious nature of the species resulting in herbarium specimens with either only male flowers, female flowers, or fruits. This could also be due to the inaccessibility of the flowering twigs during collection because they are often located in the canopy and therefore easily overlooked. Identification of *Tetrastigma* plants therefore usually has to rely on vegetative characters as was also the case in the present study since flowering or fruiting specimens of *T. loheri* were not often found during my two collecting trips in November 2016 and 2017. Sterile specimens of *T. loheri* can be distinguished from other species of *Tetrastigma* found in Cebu by a combination of features: coriaceous leaves that almost always have three leaflets (very occasionally unifoliolate) and never more than that, and young stems and leaves that are usually red (Appendix 11C). Other *Tetrastigma* species in Cebu have green young stems and leaves. *T. ellipticum* Merr., *T. harmandii* Planch., *T. scariosum* Planch., and *T.* sp. A (sensu Pelser et al., 2016b; which is an unknown species that also seems to be common in Cebu) have leaves with always more than three leaflets (five and/or seven leaflets). Although it has trifoliolate leaves like *T. loheri*, *T. papillosum* Planch. can be easily recognized by its characteristic tubercle-like epidermal protuberances on stems and petioles. Putative *T. loheri* specimens were compared with *Tetrastigma* herbarium specimens at CANU and photographs available through Co's Digital Flora of the Philippines (www.philippineplants.org; Pelser et al., 2011 onwards). The identity of some specimens was confirmed by sequencing as outlined in Chapter 3.

4.4.3 **Plant collecting**

Leaf samples from a total of 17 sampling sites were obtained from 25–47 individuals for each of the four forest areas (Figure 4.1; Appendix 12). A total of 145 individuals were collected. *Tetrastigma loheri* vines can be large and far-reaching, and multiple individuals can sometimes be found entangled. Furthermore, field observations by Barcelona and Pelser (pers. comm.) suggest that *T. loheri* might reproduce vegetatively when shoots that are resting on the soil surface develop roots and grow into shoots that eventually become disconnected from the rest of the plant. Because I aimed to sample genetically distinct individuals, I therefore took care to sample individuals that were at least 10 m away from each other to avoid unknowingly sampling the same plant twice.

Leaf samples were dried in a plastic packet with silica gel. Voucher specimens were deposited at the University of San Carlos Herbarium (CEBU) in Cebu and University of Canterbury Herbarium (CANU) in New Zealand (Appendix 12).

4.4.4 DNA extraction and development of microsatellite markers

Silica-dried leaf tissue of *T. loheri* was disrupted in a 1.5 ml microcentrifuge tube containing two metal beads using an Oscillating Mill MM400 (Retsch, GmbH, Haan, Germany). Whole genomic DNA was extracted from the pulverized leaf samples using the QIAGEN DNeasy Plant mini kit (QIAGEN, Germantown, Maryland) following the manufacturer's protocol. Extracted DNA of one individual was sent to Macrogen (South Korea) for Next Generation Sequencing (GS-FLX Titanium). This resulted in 139,776 DNA sequence reads with a total of 53,459,014 bases and an average read length of 382 bp. Using MSATCOMMANDER 1.0.8-beta (Faircloth 2008), the DNA sequence reads were screened for microsatellite loci of two to six nucleotides and M13-tagged primers were designed for them with the following settings: melting temperature: 58–62°C; GC content 30–70%, product size 100–450 base pairs (bp). The software found 2,316 microsatellite loci for which 255 primers were developed. A total of 40 primers were selected to test for successful amplification (see details below) using eight *T. loheri* specimens from a previous study (Pelser et al., 2016a). A total of 25 loci with 8–15 repeats of 2–3 nucleotide motifs with pair product size between 100–450 bp were selected for screening 145 *T. loheri* samples. Of these, three loci were excluded from subsequent analyses because they routinely displayed more than two alleles in at least of half of the total samples. A further nine loci were excluded because they were either monomorphic or failed to amplify consistently. The remaining 13 microsatellite loci were included in the analysis. The present study generated the first set of microsatellite primers for *T. loheri* (Table 4.1).

4.4.5 Microsatellite amplification and genotyping

Amplification of microsatellite loci was performed in a 4 µl PCR reaction containing the following: 1µl genomic DNA (0.76–88.39 ng/µl), 0.25 µM of untagged primer, 0.0625 µM of M13-labelled (5'-GGAAACAGCTATGACCAT-3') primer, 0.25 µM of M13 fluorescent primer (6FAM, NED, PET, or VIC), 1x of Type-it Microsatellite (QIAGEN, Germantown, Maryland) and nuclease free water to volume. A multiplex PCR assay with up to three primer combinations was performed using the following PCR settings: initial denaturation at 95°C for 15 minutes; followed by 8 cycles of 94°C for 30 seconds, 60°C for 90 seconds, 72°C for 60 seconds; 25 cycles of 94°C for 30 seconds, 52°C for 90 seconds, 72°C for 60 seconds; and a final extension at 60°C for 30 seconds. To check for successful amplification, PCR products were run on a SYBR Safe-stained gel at 100V for 30 minutes and were photographed with a Syngene G: BOXEF2 imager. Amplified products were then

analysed using the ABI 3130xL Genetic Analyzer at the University of Canterbury in New Zealand. Geneious 6.1.8 (Biomatters, Auckland, New Zealand) was used to determine fragment lengths (bp) of the amplified loci. A table of scores for the microsatellite loci and the morphology of their alleles in the chromatogram is provided in Appendices 13–14. Most DNA samples were genotyped twice. In part, this was done to determine the reproducibility of observed allele calls, but genotyping was also repeated for samples that displayed low signal strength or more than two alleles per locus. For the latter samples, genotyping PCR was repeated after increasing the annealing temperature by 2°C to 4°C in both cycles to increase primer specificity. The same approach was used for samples that displayed 'stuttering'. Data for loci for which allele calls could not be made unambiguously after repeating the genotyping analysis were recorded as missing for the affected samples.

4.4.6 **Data analysis**

A final dataset of 13 microsatellite loci (Table 4.1, Appendix 13) was used for genetic analyses of 145 individuals sampled from the four study areas in Cebu. The percentage of missing data for each locus (2.07–28.28%) and the loci with missing data (at most 4–5) for each individual in the final dataset can be found in Appendix 13.

Hardy-Weinberg Exacts Tests were used to determine if any loci deviate from Hardy-Weinberg Equilibrium. A Linkage Disequilibrium test was used to detect the presence of linkage between loci. Both tests were performed in GENEPOP (version 4.2, Raymond and Rousset, 1995) on the web (http://genepop.curtin.edu.au) using default settings. The p-values resulting from these analyses were corrected using the Benjamini and Yekutieli method for multiple comparisons (Narum, 2006). GENEPOP was also used to determine the presence and frequency of null alleles for each locus across all areas using the method of Brookfield (1996).

For each forest area, the following genetic diversity indices were computed using GenAlEx 6.503 (Peakall and Smouse, 2012): percentage of polymorphic loci, allelic richness, heterozygosity, and F_{IS} (inbreeding coefficient). A rarefaction analysis correcting for differences in sampling size between areas was implemented when calculating allelic richness in HP rare 1.0 (Kalinowski, 2005). This was done with 30 sample "genes" as the minimum sampling size. The rarefaction-corrected allelic richness was used hereinafter. An online sign test calculator (http://www.socscistatistics.com/tests/signtest/) was used to determine the statistical difference in the allelic richness between areas at $p<0.05$ using a one-tailed hypothesis. An online Spearman's Rho calculator

(http://www.socscistatistics.com/tests/spearman/) was used to determine the relationship between the total forested area size of each of the four study areas (obtained from Gonzalez et al., 1999) and their allelic richness.

The genetic connectivity among the areas was inferred by determining population genetic structure using Analysis of Molecular Variance (AMOVA), STRUCTURE, and Discriminant Analysis of Principal Components.

AMOVA was performed in GenAlEx with 999 permutations which generated F'_{ST} estimates for *T. loheri* as well as a matrix with pairwise F'_{ST} values for the four areas. F'_{ST} is a corrected value for F_{ST} which is sensitive to the genetic variation among individuals and tends to decrease when genetic variation is high (Hedrick, 2005; Meirmans and Hedrick, 2011). It is defined as the proportion of the maximum F_{ST} that can be obtained for the level of genetic variation present within a population (Hedrick, 2005; Meirmans and Hedrick, 2011).

STRUCTURE (v.2.3.2, Pritchard et al., 2000; Falush et al., 2003, 2007) is a clustering program that assigns individuals to one or more clusters following the assumptions of Hardy-Weinberg and Linkage equilibrium (Pritchard et al., 2000; Falush et al., 2003, 2007). STRUCTURE analyses were performed using the admixture model and correlated allele frequencies. Analyses were run with k-values from 1–20 with 20 iterations for each. The maximum value of k was set to 20 because 17 genetic clusters can potentially be formed from the 17 sites sampled across the four study areas. Each analysis was run using 20 iterations for 200,000 generations of which the first 100,000 were discarded as burn-in. STRUCTURE Harvester (Earl and Von Holdt, 2012) on the web (http://taylor0.biology.ucla.edu/structureHarvester/) was used to determine the value of the optimal k following the method of Evanno et al. (2005) which defines the optimal k as having the highest value of delta k. The results of the STRUCTURE analysis were compiled using the Main Pipeline of CLUMPAK -Clustering Markov Packager Across K on the web (http://clumpak.tau.ac.il).

Using the inferred optimal value for k, STRUCTURE was subsequently run again to identify migrants, using the same settings as above. The Population Information to Test for Migrants was selected as the Ancestry model with GENSBACK set to 2 to determine the ancestry of the migrants up to the last two generations. Since the migration rates among the four study areas are unknown, different values of MIGRPRIOR (0.001, 0.01, 0.05, and 0.10) were used in the analysis, as suggested by Pritchard et al. (2000).

Patterns of genetic structure were also studied using Discriminant Analysis of Principal Components (DAPC), which is a non-model genetic clustering method. The clusters

retrieved by DAPC are grouped using a k-means algorithm. In contrast to STRUCTURE analyses, DAPC does not assume that populations are in Hardy-Weinberg and linkage equilibrium (Jombart et al. 2010), this approach is therefore often used in studies of populations that are not in Hardy-Weinberg equilibrium such as those of crops that are highly inbred and experience non-random mating (e.g., Matos et al., 2013; Filippi et al., 2015; Campoy et al., 2016). DAPC is similar to Principal Component Analysis (PCA) in that they are both multivariate methods that are capable of summarising variation into a reduced number of dimensions (Jombart et al. 2010). However, unlike PCA, DAPC can assess the number of clusters using k-means clustering. Furthermore, DAPC has a better way of visualizing the variation between clusters by maximizing the between-group differences while minimizing the within-group differences (Jombart et al., 2010). In the present study, DAPC was done in R-studio using the adegenet package (version 2.1.1; Jombart, 2008). The optimal value of k was determined using the *find.cluster* function, exploring k values ranging from k from 1 to 20 (Jombart and Collins, 2015). The k value with the lowest Bayesian Information Criterion (BIC) value (Appendix 15) was chosen as the optimal k. Using this value, the DAPC analysis was performed using the function *dapc*. The optimal number of retained principal components for the analysis was 21 (Appendix 16) which was determined by using the *optim.a.score* function. A scatterplot of individuals was generated using the *scatter* function.

The genetic diversity analyses outlined above identified a significant departure from Hardy-Weinberg equilibrium in each study area as a result of an excess of homozygotes. This can sometimes be attributed to sampling of at least two genetically distinct populations (i.e. the presence of genetic substructure within study areas): the Wahlund effect (Garnier-Géré and Chikhi, 2013). To check for the presence of genetic substructure within each study area, a separate DAPC was carried out for each area. The optimal values of k were determined using the maximum number of sampling sites per study area as the upper limit of k to be tested. For example, CCPL has four sampling sites, hence, the range of k tested for this area was from one to four.

The relationship between genetic distance and geographic distance was determined to investigate if population genetic structure can be explained by distance. The Mantel test (1967) is a statistical test used to determine the correlation between two matrices. This test of isolation by distance was employed by determining the correlation between the logarithmic geographic distance and the Nei genetic distance in GenAlEx with 999 permutations.

Table 4.1. Features of the microsatellite primers developed for Tetrastigma loheri*. Annealing temperatures were 60°C and 52°C in the first and second PCR cycle respectively. The M13 nucleotide sequence is underlined.*

Locus	Primer sequence (5'– 3')	Allele size (bp)	Repeat motif
LOH412	F: GGAAACAGCTATGACCATCCCACACTCTTCTCATGCC	229–316	$(AAT)_{14}$
	R: TCTCGAAAGTCAGGAAATGGC		
LOH505	F: GGAAACAGCTATGACCATCCTCTAAGTGCCTC	389–467	$(AAT)_{14}$
	R: CTTGCCCACAGTGCCTTTAG		
LOH663	F: GGAAACAGCTATGACCATGAACCAAACACGGCCTAAGG	288–336	$(AT)_{13}$
	R: TGCAGCTTGGTCAGTTATCTC		
LOH684	F: AGGGCATCAGATCAGACAGAC	341–347	$(AAC)_{8}$
	R: GGAAACAGCTATGACCATCCATCTCCTACCTCGCG		
LOH688	F: CTCTCGTCGCGTAAACCAAC	308–353	$(AAT)_{11}$
	R: GGAAACAGCTATGACCATCAACCCTACTGTGACCGC		
LOH749	F: GGAAACAGCTATGACCATGACAGTGCTTCCAACCAC	207–277	$(AT)_{11}$
	R: GCCACGCTCATACTCACAAG		
LOH865	F: GGAAACAGCTATGACCATGCGATGATGTTGTCCTGAGG	217–253	$(AT)_{12}$
	R: ATGTATTGTCGGGTCCCACG		
LOH868	F: GGAAACAGCTATGACCATTGTAGCGCCCTAATTCCCG	143–213	$(AT)_{13}$
	R: GGCAACAAACTAGCCAGGTC		
LOH931	F: TCGACGATCCAATGCAATCG	205–263	$(AT)_{11}$
	R: GGAAACAGCTATGACCATCAACCATCGATTTAACCACCAG		
LOH936	F: CACACCTGATTCTTGGCTCTG	197–233	$(AC)_{9}$
	R: GGAAACAGCTATGACCATACCCTCACCATAAAGAGTGTG		
LOH1248	F: AGATGAAGGTTTGCTGCTCG	187–209	$(AG)_{9}$
	R: GGAAACAGCTATGACCATTCTTCAGGTGCATCAGGATC		
LOH1497	F: TGGTAGGTGAATGGCATTGG	190–218	$(AG)_{9}$
	R: GGAAACAGCTATGACCATCACCGCCACTTTCTTCTCC		
LOH1512	F: GGAAACAGCTATGACCATCCCTAACCAAATTCAGCTCACC	141–189	$(AC)_{11}$
	R: CCGCTTATCTTTGGTCATTGC		

4.5 RESULTS

4.5.1 Data quality and assumptions

A significant departure from Hardy-Weinberg equilibrium (HWE) due to a deficit of heterozygotes was found for nine of the 13 microsatellite loci (69%) after B-Y correction ($p<0.011$). Deviation from HWE was not found in LOH 936, LOH 684, LOH 1497, and LOH 688 in all four areas. A highly significant departure from the HWE was reported by GENEPOP (no p-values provided) in all four areas. Linkage disequilibrium (LD) was significant after B-Y correction ($p<0.008$) in only 14 of 312 (4%) pairwise values between loci across the four areas. Pairs of loci in significant linkage disequilibrium were only found in CCPL and Argao. (Table 4.2). Nine of these were from CCPL and five were from Argao. One of these 14 pairs of loci was found in both CCPL and Argao. These 14 pairs of loci were not considered evidence of linkage disequilibrium in the present study because they were not found across all four areas. The results of the GENEPOP analysis suggested that eight of the 13 loci (62%) display null alleles across the four areas.

Table 4.2. List of loci pairs significantly deviating from linkage equilibrium ($p<0.008$ after B-Y correction) in CCPL and Argao. Loci pairs with asterisk occur in both areas.

CCPL	Argao
LOH 931/ LOH 749	LOH 931/ LOH 936
LOH 749/ LOH 1497	LOH 663/ LOH 868
LOH 868/ LOH 1497	LOH 868/ LOH 688
LOH 412/ LOH 688	*LOH 749/ LOH 865
LOH 749/ LOH 688	LOH 868/ LOH 1512
LOH 688/ LOH 505	
*LOH 749/ LOH 865	
LOH 412/ LOH 1512	
LOH 1497/ LOH 1512	

4.5.2 Genetic diversity of *T. loheri* populations

A total of 226 different alleles from 13 loci were found. All loci were polymorphic with an average of 17.38 alleles per locus. Allelic richness (corrected for sample size) in each area ranged from 6.99 to 8.53, with Dalaguete having the lowest and CCPL the highest richness (Table 4.3). CCPL has significantly higher allelic richness than Dalaguete (p=0.006. Likewise, a higher allelic richness was found in Argao than in Dalaguete (p=0.042). The correlation between forest cover and allelic richness was not statistically significant (R= –0.8, p=0.2; Table 4.3).

Observed heterozygosity was much lower than the expected heterozygosity across all four forest areas (average observed heterozygosity =0.51, the average expected heterozygosity=0.72) indicating the presence of excess homozygotes (Table 4.3). The F_{IS} in each area ranged between 0.23 and 0.31 (average 0.26; Table 4.3).

4.5.3 Genetic structure among *T. loheri* populations

Patterns of genetic structure were studied to infer genetic connectivity among the four areas and were investigated using three approaches. Results of the AMOVA revealed that 7% of the variation was found among populations and 38% was among individuals (Table 4.4). The F'_{ST} (0.307) was significant at p=0.001 which indicates a relatively high level of genetic differentiation (Wright 1978) among the four areas. Pairwise comparison of F'_{ST} of the four sites showed significant genetic differentiation for all pairs (Table 4.5) indicating that each site is genetically different from each other. The highest pairwise F'_{ST} was found between CCPL and Dalaguete. The lowest pairwise F'_{ST} was found between Argao and Dalaguete.

STRUCTURE identified four genetic clusters following the method of Evanno et al. (2005) (Appendix 17). The STRUCTURE plot of membership showed that most individuals have high probabilities of belonging to the same genetic cluster as other specimens from the same area (Figure 4.2). However, the plot also showed the presence of a number of admixed individuals in the four areas suggesting the presence of migrants. The highest number of migrants was detected when the migration rate was set at 0.1. Four admixed individuals (*Obico 409*, *Obico 427*, *Obico 458*, and *Obico 484*) were identified as migrants or descendants of migrants (Table 4.6; Appendix 18). *Obico 427* was determined to be a recent migrant from Dalaguete to Argao. *Obico 409, 458*, and *484* were identified as second-generation migrants. *Obico 409* in Argao had one or more grandparents originally from Dalaguete. *Obico 458* in Argao had one or more grandparents originally from CCPL or Dalaguete. *Obico 484* in Dalaguete had one or more grandparents from CCPL.

The DAPC identified four genetic clusters as revealed by the lowest corresponding BIC value (Appendix 15). This was the same number as detected by STRUCTURE. The scatterplot showed that individuals were grouped into four clusters which largely corresponded to the four areas (Figure 4.3), although the Argao and Dalaguete clusters were partially overlapping. To determine genetic substructure, a DAPC was carried out for each of the four areas. The results of these DAPC showed that none of the four areas has genetic substructure as revealed by BIC values of 1 (not shown), which suggests the presence of only one cluster per area.

A positive, relatively weak, but significant correlation between genetic and geographic distances (R= 0.269, p=0.001) was detected by a Mantel test (Appendix 19).

Table 4.3. Standard genetic diversity indices observed at 13 microsatellite loci for populations of Tetrastigma loheri *in four areas in Cebu. Column labels: N-Number of samples, Size (ha)- forest cover in hectare from Gonzalez et al. (1999), P-percentage of polymorphic loci, Na-allelic richness corrected for sample size, Ho-observed heterozygosity, He-expected heterozygosity, and F_{IS}-inbreeding coefficient. Standard error values are in parenthesis. Means with the same superscript in the Na column are not significant at p=0.05.*

Areas	**N**	**Size (ha)**	**P**	**Na**	**Ho**	**He**	**F_{IS}**
CCPL	47	420	100%	8.53[a]	0.54 (0.07)	0.72 (0.06)	0.23 (0.08)
Argao	46	507	100%	8.21[a]	0.53 (0.06)	0.71 (0.06)	0.24 (0.07)
Dalaguete	25	654	100%	6.99[b]	0.47 (0.05)	0.69 (0.05)	0.31 (0.07)
Alcoy	27	1,036	100%	7.78[ab]	0.51 (0.06)	0.73 (0.05)	0.27 (0.09)
Average				7.8775	0.51 (0.03)	0.72 (0.03)	0.26 (0.04)

Table 4.4. Partitioning of variance and F-statistics derived from the Analysis of Molecular Variance for populations of Tetrastigma loheri *in four areas in Cebu. Column labels: Degrees of freedom (df), Sum of squares (SS), Mean square (MS), Estimated variance (Est. Var.), Percentage variance (%).*

Source	**df**	**SS**	**MS**	**Est. Var.**	**%**
Among Populations	3	103.228	34.409	0.389	7%
Among Individuals	141	983.062	6.972	2.012	38%
Within Individuals	145	427.500	2.948	2.948	55%
Total	289	1513.790		5.349	100%
F-Statistics	**Value**	**P(rand >= data)**			
Fst	0.073	0.001			
Fis	0.406	0.001			
Fit	0.449	0.001			
F'st	0.307				

Table 4.5. Pairwise F'st values for Tetrastigma loheri *between the four areas. All values are significant at p=0.02 after B-Y method correction (Narum, 2006).*

CCPL	Argao	Dalaguete	
			CCPL
0.259			Argao
0.387	0.243		Dalaguete
0.327	0.335	0.346	Alcoy

Table 4.6. Individuals of Tetrastigma loheri *s.l. that were detected as migrants or descendants of migrants when migration prior was set at 0.10. Column labels: Specimen no. refers to the collecting number of the specimen with the label of the individual in the STRUCTURE plot enclosed in parenthesis (Appendix 18). Population refers to the four forested areas in Cebu where the sampling origin of an individual is indicated as 'origin' in parenthesis. Generation 0 indicates the probability of an individual being a migrant. Generation 1 indicates the probability of an individual belonging to a first-generation migrant. Generation 2 indicates the probability of an individual belonging to a second-generation migrant.*

		Probability of being from assumed population		
Specimen no.	Population	Generation 0	Generation 1	Generation 2
Obico 409 (51)	Argao (origin)	0.45		
	CCPL	0.002	0	0.018
	Dalaguete	0.138	0.014	0.356
	Alcoy	0.001	0	0.003
Obico 427 (62)	Argao (origin)	0.135		
	CCPL	0	0.001	0.004
	Dalaguete	0.731	0.08	0.044
	Alcoy	0	0	0.003
Obico 458 (90)	Argao (origin)	0.44		
	CCPL	0.003	0.037	0.166
	Dalaguete	0.021	0.083	0.163
	Alcoy	0.004	0.009	0.075
Obico 484 (103)	Dalaguete (origin)	0.424		
	CCPL	0.075	0.044	0.334
	Argao	0.028	0.009	0.053
	Alcoy	0.004	0.001	0.014

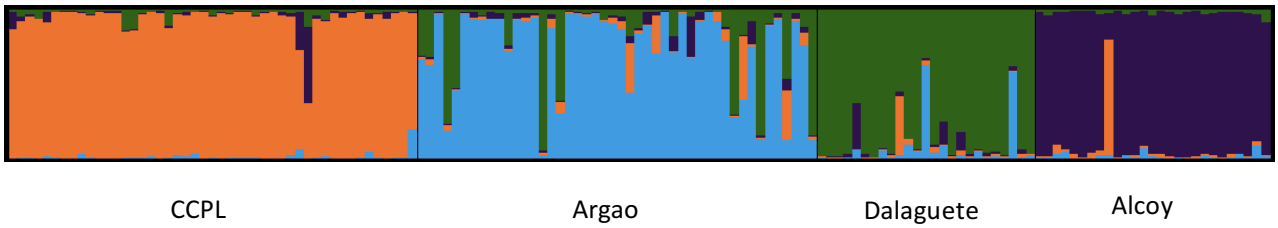


Figure 4.2. STRUCTURE plot (k=4) of membership probabilities of Tetrastigma loheri *individuals in the four study areas. Each bar represents one individual and the proportion of each colour in each bar shows membership probabilities.*

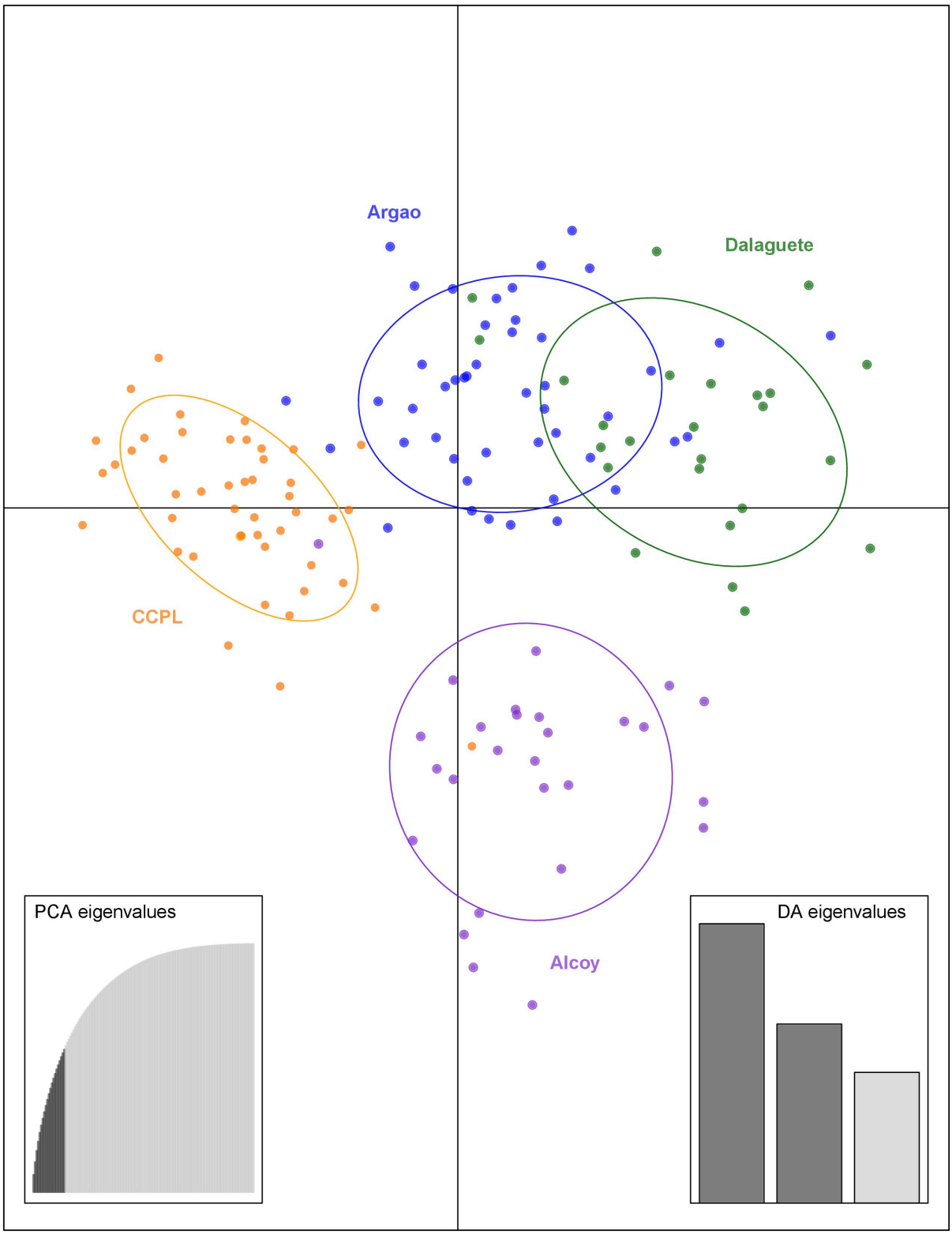


Figure 4.3. Discriminant Analysis of Principal Components (Jombart et al., 2010) scatterplot of 145 individuals of Tetrastigma loheri *grouped into 4 genetic clusters. The different colours refer to the four forest areas. Left-hand inset shows proportion of PCA eigenvalues retained in the analysis. Right-hand inset shows the first three discriminant analysis eigenvalues used in the analysis. Circles represent inertia ellipses encircling 2/3 of the points in each cluster.*

4.6 DISCUSSION

Almost all original forest cover on Cebu has been destroyed in recent centuries and only a few forested areas remain (Figure 4.1; Gonzalez et al., 1999). These areas support most of the remaining native biodiversity of the island that requires natural forest habitat (Gonzalez et al., 1999; Paguntalan et al., 2015). Because of the small size of these forested areas (i.e. up to 1,036 ha) and the large distances between several of them (Figure 4.1; Gonzalez et al., 1999; Paguntalan et al., 2015), many plant populations in these forested areas are likely to be smaller and more isolated from each other than they were prior to the destruction and fragmentation of Cebu's natural forests. Although population genetic theory predicts that this would have resulted in a loss of genetic diversity (Ellstrand & Elam, 1993), presently, there are to my knowledge no publicly available genetic data that provide an indication of the patterns of remaining genetic diversity and connectivity among forested areas in Cebu. However, this information is important in conservation management of the remaining forested areas that aims to maintain genetic diversity for the long-term persistence of species. In this study, a population genetic study of *Tetrastigma loheri* in the four largest forested areas of Cebu was carried out as a first step towards understanding the patterns of genetic diversity and connectivity among the remaining forested areas of Cebu.

4.6.1 Genetic connectivity

The results of the AMOVA (F'st= 0.307, p=0.001) indicate the presence of genetic differentiation between all four areas (Table 4.4; all pairwise F'st values statistically significant at p= 0.02), suggesting low levels of genetic connectivity between them. The presence of genetic differentiation is also indicated in the results of both the STRUCTURE analysis and the DAPC, which identified four genetic clusters that largely corresponded to the four areas (Figures 4.2 and 4.3). These results are in agreement with the prediction by population genetic theory that habitat fragmentation reduces genetic connectivity (Lowe et al., 2005).

Genetic connectivity between plant populations is facilitated by the transport of pollen and seed between them (Sork & Smouse, 2006). Vectors of pollen and seed can be non-biological like wind or biological such as animals. Studies aimed at identifying the pollinators and seed dispersers of *T. loheri* have, to my knowledge, not been performed. However, data from other *Tetrastigma* species (e.g., *Tetrastigma diepenhorstii*; Zakaria et al., 2017) and the morphology of *T. loheri*'s reproductive structures indicate that the species is most likely pollinated by insects and that its seeds are most likely dispersed by birds. A

Mantel test shows a significant positive correlation (R= 0.269, p=0.001) between genetic and geographic distance (Appendix 19) suggesting isolation by distance. This might indicate that the ability of the pollinators or seed dispersers to move between areas is negatively correlated with the distances between the four areas, resulting in lower genetic connectivity between them. The distance between CCPL and the group of three southern forest areas (Alcoy, Argao and Dalaguete) is about 60–90 km. The three southern forest areas are about 10–25 km apart from each other. These distances might be large enough to prevent travel of pollinators or seed dispersers between the four areas. Future studies might investigate the reproductive ecology of *T. loheri* to study the behaviour of its pollinators and seed dispersers and understand how the distances between the four areas affect the dispersal of the pollen and seed of *T. loheri*.

Argao and Dalaguete are the forest areas with the best genetic connectivity between them, as inferred from their lowest pairwise F'_{ST} of all pairs of areas (i.e. 0.243; Table 4.5). These results are corroborated by the results from STRUCTURE and the DAPC. The presence of several admixed Argao individuals that have high membership probabilities to Dalaguete (Figure 4.2), and vice versa, suggests some migration between these two forest areas. Indeed, three of the four migration events identified by STRUCTURE involved individuals that potentially migrated between Argao and Dalaguete, supporting higher connectivity between them relative to other forest areas. This level of genetic connectivity between Argao and Dalaguete can also be inferred from the overlapping clusters of Argao and Dalaguete in the DAPC plot (Figure 4.3). This level of genetic connectivity between Argao and Dalaguete is not seen between Dalaguete and Alcoy although the distances among Argao, Dalaguete, and Alcoy are similar (i.e., c. 10–15 km). In fact, the pairwise F'st value between Dalaguete and Alcoy was the second highest of all pairs of areas (i.e. 0.346), which was unexpected considering their proximity.

Further, it is possible that the isolation by distance results may be an artefact of spatial autocorrelation, which can bias the result of a Mantel test leading to false positives (Meirmans, 2012). Individuals that are closer together are expected to be more genetically similar than individuals that are further apart. This can result in spatial autocorrelation (Meirmans, 2012). One criticism on the use of Mantel test to test isolation by distance is that it does not distinguish between the patterns resulting from spatial autocorrelation and isolation by distance (Meirmans, 2012). The possibility of the existence of spatial

autocorrelation might serve as a caveat in using isolation by distance to explain the observed patterns of genetic differentiation in this study.

Regardless, geographical distance seems unlikely to be the sole explanation for the observed patterns of genetic differentiation in Cebu. Although there are no natural barriers such as tall mountains, deep valleys or bodies of water between the forested areas, poor quality habitat or inhospitable terrain between them is evident and might affect foraging behaviour, thus, affecting the movement of pollinators and seed dispersers between populations (Aguilar et al., 2008; Vranckx et al., 2012). For example, insect pollinators have been shown to remain within habitat fragments (Goverde et al., 2002), thus, reducing pollen movement. Similarly, forest fragmentation leaves forest gaps that can act as barriers to the movement of birds, particularly forest specialists and understory species (Harris & Reed, 2002). Furthermore, habitat isolation can also affect the availability of pollinators and seed dispersers. For instance, pollinator abundance and diversity have been found to decline in areas that have experienced habitat fragmentation (Rathcke & Jules, 1993) and bird dispersers have been shown to have disappeared from small residual forest fragments or to be less common than in continuous forests (Uriarte et al., 2011). The four forest areas of this study are all relatively small (Gonzalez et al., 1999) and the landscape between the four forest areas is mostly a mosaic of roads, farmlands, clearings, and residential areas (Appendices 20–22), creating various levels of disturbance that might impede the movement of insects and birds between the areas (Fischer & Lindenmayer, 2007; Barr et al., 2015). Determining the effect of isolation by environment through landscape genetic studies and studies aimed at discovering how the intervening areas affect the behaviour of pollinators and seed dispersers (e.g., Wang & Bradburd, 2014; Manthey & Moyle, 2015) are relevant topics to investigate in the future.

4.6.2 **Inbreeding in *T. loheri* populations**

The excess of homozygotes in all of the four areas resulted in a significant departure from Hardy-Weinberg equilibrium. Observing excess homozygosity in populations can be attributed to several factors, such as the presence of null alleles, the Wahlund effect, or inbreeding (Selkoe & Toonen, 2006). Because of the relatively high proportion of loci identified as having null alleles (62%), it is more likely that the high level of homozygosity observed in this study is a consequence of biological factors such as inbreeding or the Wahlund effect (Rossetto et al., 1999; Dakin and Avise, 2004; Dewoody, 2006). To determine if the Wahlund effect (Sinnock, 1975) might explain the results, DAPC analyses

were carried out to test for the presence of genetic substructuring. Genetic substructure indicates the presence of genetically distinct subpopulations. When genetic data from two or more genetically distinct subpopulations are combined and their genetic diversity are estimated together as one population, this results in reduced heterozygosity. No evidence of genetic substructuring was found in any of the four areas (data not shown), indicating that it is not likely that the Wahlund effect explains the observed excess of homozygotes. The most plausible reason for the high level of homozygosity observed in this study is therefore inbreeding.

4.6.3 **High levels of inbreeding with no evidence of low genetic diversity**

Tetrastigma loheri populations showed a high level of inbreeding (F_{IS} =0.23-0.31; Table 4.3) compared to other plant species with similar life history traits (F_{IS} = 0.09-0.15; Table 4.7) as recorded by Nybom (2004) and Ballesteros-Mejia et al. (2016). Despite high levels of inbreeding, no evidence of low genetic diversity was found in *T. loheri* populations. The expected heterozygosity (He=0.72) of *T. loheri* in Cebu falls within the range of values observed using microsatellite markers from species with similar life history traits (He=0.62–0.80, as summarized by Nybom (2004) and Ballesteros-Mejia et al. (2016); Table 4.7). In addition, the estimated allelic richness of *T. loheri* is higher (7.88) than that of other plants with a similar life history (4.18–6.00; Ballesteros-Mejia et al. (2016); Table 4.7). Overall, these data indicate that the genetic diversity in *T. loheri* populations is not substantially lower than that of biologically similar plant species.

Habitat fragmentation can result in inbreeding and genetic diversity loss due to genetic drift, in part due to a lack of genetic connectivity between fragments (Young et al., 1996). Consequently, mating of plants in isolated fragmented forests may occur between related individuals, especially if these isolated populations are small (Aguilar et al., 2008; Ellstrand & Elam, 1993; Young et al., 1996). However, genetic diversity loss might not manifest in fragmented populations with relatively large historical effective population sizes (e.g., Turchetto et al., 2018). Effective population size is defined as "the size of an idealized population that would give rise to the same variance of gene frequency, or rate of inbreeding as in the actual population under consideration" (Frankham, 1995). Populations with small effective population sizes are susceptible to loss of genetic diversity through genetic drift (Ellstrand & Elam, 1993) and suffer most when their populations are reduced (Templeton et al., 1990). Populations with large effective population sizes, therefore, are

not as vulnerable to genetic diversity loss as populations with small effective population sizes.

Furthermore, genetic diversity loss might not manifest in fragmented populations if the fragmentation event was relatively recent and the species are long-lived. In a habitat that has experienced a recent fragmentation event, individuals of long-lived species (e.g., woody trees) might represent remnant populations of the pre-fragmented forest (Kramer et al., 2008). These long-lived species are less likely to lose genetic diversity than herbaceous plants because of their long generation times (Young et al., 1996; Lowe et al., 2005; Kramer et al., 2008). However, a meta-analysis has shown that some trees and shrubs are as susceptible as herbaceous species to the negative effects of habitat fragmentation (Vranckx et al., 2012). Considering *T. loheri* is a commonly encountered plant in the forested areas in Cebu, it is possible that it had a large population size in the past. Because the process of the deforestation of Cebu only started a few centuries ago (Bankoff, 2007), the fragmentation that resulted in the remaining small forested areas, might have been relatively recent (Kramer et al. 2008). It is unknown if *T. loheri* is a long-lived species. Although woody, the main stem of *T. loheri* has not been observed to grow a relatively thick diameter (Pelser pers. comm.). This therefore suggests that it is not likely that *T. loheri* are long-lived plants. However, plants might be longer lived than their stem diameter suggests, because *T. loheri* might be capable of reproducing vegetatively (Barcelona and Pelser pers. obs.). New individuals could arise from stems that contact the soil surface and develop roots, eventually disconnecting from the mother plant as new shoots develop. Plants that reproduce clonally through such vegetative means may be less vulnerable to the genetic effects of fragmentation, because their genotypes are effectively long-lived (Honnay & Bossuyt, 2005). However, the presence and prevalence of asexual reproduction in *T. loheri* remains to be further studied.

In conclusion, the absence of evidence of low genetic diversity despite inbreeding in *T. loheri* populations in Cebu might be attributed to a large historical effective population size, the relative recent forest fragmentation in Cebu, and the possibility that *T. loheri* is a long-lived liana.

Table 4.7. Comparison of expected heterozygosity, allelic richness, and inbreeding coefficients recorded in the present study, and two review papers summarising the expected values for different life histories: Nybom (2004) and Ballesteros-Mejia et al. (2016).

	Expected heterozygosity			**Allelic richness**		**Inbreeding coeffieicent (F_{IS})**	
Life history traits	**Nybom (2004)**	**Ballesteros-Mejia et al. (2016)**	**This study**	**Ballesteros-Mejia et al. (2016)**	**This study**	**Ballesteros-Mejia et al. (2016)**	**This study**
Widespread	0.62	0.752±0.131	0.72	6.009±3.567	7.88	0.117±0.096	0.26
Outcrossing	0.65	0.767±0.122	0.72	Not included	NA	0.115±0.095	0.26
Dioecious	Not included	0.622±0.000	0.72	4.178±2.146	7.88	0.154±0.217	0.26
Seeds ingested	0.73	Not included	0.72	Not included	NA	Not included	NA
Bird dispersal	Not included	Not included	NA	4.973±3.021	7.88	0.148±0.126	0.26
Bee and wasp pollination	Not included	0.803±0.105	0.72	4.857±2.890	7.88	0.117±0.106	0.26
Moth and butterfly pollination	Not included	Not included	NA	5.241±3.013	7.88	0.087±0.078	0.26

4.6.4 **Conservation implications and recommendations**

This study provides a first glimpse of patterns of genetic diversity and connectivity of fragmented populations of forest species in Cebu. Its results suggest that the negative consequences of habitat fragmentation, such as inbreeding and low genetic connectivity, are likely to be greater for plant species that are less common than *T. loheri* and that have smaller population sizes or more restricted distributions (e.g. *Cynometra cebuensis* and *Cinnamomum cebuense*; Department of Environment and Natural Resources Administrative Order, 2017)

Conservation genetics is an emerging field in Cebu, in particular, and the Philippines, in general. Because conservation genetic data for endangered plant species in Cebu are currently lacking, using our data to make recommendations for conservation management should follow a precautionary approach. Under the assumption that other plant species than *T. loheri* are also experiencing low genetic connectivity, and might likewise be experiencing inbreeding, and perhaps even a loss of genetic diversity, we recommend the establishment of ecological corridors to increase genetic connectivity between the four areas to reduce the risk of inbreeding and loss of genetic diversity. Corridors have been shown to increase genetic connectivity between isolated populations separated by forest fragmentation by allowing

movement of pollinators and seed dispersers such as insects and birds (Tewksbury et al., 2002; Levey et al., 2005; Damschen et al., 2006). Since CCPL is relatively far from the three southern areas (Figure 4.1), it would be more practical to establish corridors with the two other forested areas in the northern part of the island, i.e. Tuburan and Catmon. The three southern areas are relatively close to each other and establishing genetic corridors of forest habitat between them or with two other southwestern forested areas, Alegria and Malabuyoc (Figure 4.1), is feasible and is therefore recommended.

Nug-as forest in Alcoy and Boljoon, and the remaining fragmented forests in Argao and Dalaguete should also be awarded formal protective status to prevent further destruction, fragmentation, and degradation. However, protective status is not likely to be effective without community support. This is evident in the CCPL, which is formally a protected area, but is experiencing ongoing decline. In contrast, Nug-as forest is not a formally protected area and, as such, does not have a Protected Area Management Board. However, it is well managed by a local Peoples Organisation supported by a non-government organisation (Philippines Biodiversity Conservation Foundation, Inc.) and has witnessed habitat regeneration in the recent decade.

CHAPTER 5: Summary and research outlook

Conservation genetic data inform conservation management that aims to maintain genetic diversity for the long-term persistence of species (Jamieson et al., 2008). For Cebu, a large Philippine island that has lost nearly all of its forest cover (Gonzalez et al., 1999; Jakosalem et al., 2013; Paguntalan et al., 2015), published population-level genetic data for plant species are not available. The effects of habitat fragmentation on the genetic diversity of forest plant species in Cebu are therefore currently unknown. To address this lack of knowledge, I studied patterns of genetic diversity and genetic connectivity for the four largest remaining forest areas in Cebu (Alcoy, Argao, Dalaguete, and the Central Cebu Protected Landscape (CCPL)) using microsatellite data of *Tetrastigma loheri* Gagnep. (Vitaceae), a commonly encountered forest vine species. However, because previous research (Pelser et al., 2016b) indicated that this species might be a member of a species complex, I first studied the species delimitation of *T. loheri* and other putative members of the complex in the Philippines (i.e. the *T. loheri* s.l. complex) using unsupervised and supervised clustering analyses of a dataset of vegetative morphological characters, as well as species delimitation models on a DNA sequence phylogeny of the complex. In this chapter, I outline the major findings of my PhD study and present some areas for further research.

5.1 MORPHOLOGICAL DIVERSITY IN LEAF SHAPE AND OTHER VEGETATIVE CHARACTERS SUGGESTS THAT THE *TETRASTIGMA LOHERI* S.L. COMPLEX IS NOT COMPOSED OF MORE THAN ONE SPECIES (CHAPTER 2).

Previous research (e.g., Pelser et al., 2016b) suggested that *T. loheri* forms a species complex with the morphologically similar *T. philippinense* Merr., *T. stenophyllum* Merr., and Philippine representatives of *T. trifoliolatum* Merr. and *T. diepenhorstii* (Miq.) Latiff. In Chapter 2, as one approach to resolving the delimitation of the *T. loheri* s.l. complex, I explored data that can be obtained from vegetative characters, which are the most accessible features of *T. loheri* s.l. for taxonomic study since their reproductive parts are poorly known. Using *T. loheri* s.l. specimens collected from different provinces across the Philippines (including type localities), a leaf shape diversity study using geometric morphometric approaches and a morphological study of other vegetative characters were conducted. This was done to find morphological discontinuities among groups of specimens, which could be interpreted as evidence for the presence of more than one species within *T. loheri* s.l using

morphological diagnosability as an indication of lineage separation under the unified species concept (De Queiroz, 2007). Unfortunately, *T. diepenhorstii* could not be included in this study, because specimens of this species from the Philippines and elsewhere were not available for study.

The results of unsupervised clustering analyses of the geometric morphometric and other morphological data revealed the absence of distinct morphological clusters of individuals. They therefore do not provide conclusive evidence for the existence of more than one species within *T. loheri* s.l. in the Philippines. Under the assumption that the vegetative characters that were studied are suitable characters for species delimitation in *T. loheri* s.l., these results imply that *T. loheri, T. philippinense*, *T. stenophyllum*, and *T. trifoliolatum* are conspecific. Of these, the name *T. loheri* Gagnep. has priority (Turland et al. 2018). However, taxonomic changes were not proposed, because it remains possible that more than one species exists in the *T. loheri* s.l. complex. The possibility that these are phylogenetically distinct, but morphologically cryptic and only diagnosable following supervised morphometric analyses was explored in Chapter 3.

5.2 SPECIES DELIMITATION MODELLING DOES NOT PROVIDE EVIDENCE FOR THE EXISTENCE OF MORE THAN ONE SPECIES WITHIN THE *TETRASTIGMA LOHERI* S.L. SPECIES COMPLEX IN THE PHILIPPINES (CHAPTER 3)

In Chapter 3, I continued to study the delimitation of the *T. loheri* s.l. complex by determining if *T. loheri* s.l. consists of groups of specimens that are monophyletic and morphologically distinct. This would indicate the presence of multiple species under a unified species concept (De Queiroz, 2007) using both monophyly and diagnosability as evidence of lineage separation. Two species delimitation models, the Generalized Mixed Yule Coalescent (GMYC; Pons et al., 2006; Monaghan et al., 2009; Fujisawa & Barraclough, 2013) and Poisson Tree Processes (PTP; Zhang et al., 2013) models, were employed to find putative species that form monophyletic groups in a Bayesian Inference phylogeny obtained from a combination of DNA sequence data from five plastid and two nuclear DNA regions. A supervised clustering method using Random Forest analysis was subsequently used to identify vegetative characters that are diagnostic for these monophyletic groups.

GMYC and PTP identified several monophyletic groups as putative species. However, these are statistically poorly-supported and Random Forest analyses did not result

in the identification of vegetative characters that support these monophyletic groups. The results of the morphometric and phylogenetic analyses presented in Chapters 2 and 3 therefore do not provide conclusive evidence in support of recognizing more than one species within *T. loheri* s.l. in the Philippines. The nested position of accessions of *T. diepenhorstii* among other *T. loheri* s.l. accessions suggests that, in addition to *T. philippinense, T. stenophyllum,* and *T. trifoliolatum*, this species might also be conspecific with *T. loheri*. However, type specimens or specimens collected from the type locality of *T. diepenhorstii* were not included in this study. This would be important for nomenclatural reasons (the name *T. diepenhorstii* (Miq.) Planch. has priority over *T. loheri* Gagnep. if both are indeed synonyms) and necessary in future taxonomic studies aimed at resolving the *T. loheri* s.l. complex that would encompass the entire distribution area of the complex instead of only the Philippines. Ideally, these future studies would also include representatives of other taxa that are morphologically similar to those included in this study, but that do not occur in the Philippines and were therefore not studied as part of my project (e.g., *T. curtisii* (Ridl.) Suess.; Zakaria et al., 2016). Such future research should also aim to obtain a better understanding of the reproductive morphology of the complex, so that these characters can be used to determine if it consists of species that are cryptic in their vegetative morphology, but diagnosably distinct in reproductive features. This research would only be feasible, however, if a targeted effort is made to find, collect and preserve flowering and fruiting *Tetrastigma* specimens, because only very few of these are available and this was identified as a significant barrier in the present study.

Although not recognised as a distinct species in this study because of a lack of morphological support, Group H is best supported as taxonomically distinct at the species level among the groups delimited by GMYC and PTP. Group H forms a strongly supported clade and is more distantly related to the other *T. loheri* s.l. groups than they are to each other. Furthermore, Group H contains specimens that are sympatric with members of other groups and is therefore potentially reproductively isolated from them. In addition, Group H is the group for which the lowest Random Forest classification errors were obtained in the GMYC and PTP confusion matrices, suggesting that there might be some morphological features in which it is different from the remainder. Group H is therefore a primary candidate for further species delimitation studies.

The present conservation genetics project as well as a previous study that aimed to obtain a better understanding of the host specificity of Philippine *Rafflesia* R.Br., which only parasitize *Tetrastigma* species (Pelser et al., 2016b), clearly demonstrate the need for a

comprehensive taxonomic revision of *Tetrastigma* throughout its distribution area. These two studies addressed quite different biological questions, yet were both complicated by the absence of a modern taxonomic framework for this genus that provides clarity about the number of species of *Tetrastigma*, their delimitation, their distribution, and the scientific names that should be used for them. Considering the scale and complexity of revising *Tetrastigma*, a genus of c. 95 species with a distribution area that ranges from the Himalayas to southeastern Australia (Chen et al., 2001), such a comprehensive taxonomic treatment might not be available in the near future. However, progress towards it may be achieved through collecting specimens of species that have not yet been included in systematic studies, particularly at the type localities of the most poorly known species, and using molecular phylogenetic approaches to determine their relationships relative to other *Tetrastigma* species.

5.3 NO EVIDENCE OF LOW GENETIC DIVERSITY DESPITE HIGH LEVELS OF INBREEDING AND POOR GENETIC CONNECTIVITY AMONG *TETRASTIGMA LOHERI* POPULATIONS IN REMAINING FOREST AREAS IN CEBU, PHILIPPINES (CHAPTER 4)

Although not entirely conclusive, the results of the taxonomic studies presented in this thesis indicate that the *T. loheri* s.l. complex is only composed of one species in the Philippines. Even if this conclusion proves incorrect following future research in which reproductive characters are studied, the results of the present study still suggest that the *T. loheri* specimens that were sampled from Cebu for the conservation genetic component of my thesis research are conspecific, because specimens from all four areas that were the target of my study (Alcoy, Argao, Dalaguete, and CCPL) were resolved as part of the same GMYC and PTP groups. This means that the *T. loheri* microsatellite dataset that I compiled can be used for studying patterns of genetic diversity and genetic connectivity among these remaining forested areas.

The results of population genetic analyses of the microsatellite dataset do not reveal evidence of low genetic diversity, despite suggesting a relatively high level of inbreeding in the four forested areas. Furthermore, low levels of genetic connectivity were evident among these areas, as inferred from identifying significant genetic differentiation among them. These findings suggest that the negative consequences of habitat fragmentation, such as inbreeding and low genetic connectivity, are likely to be greater for plant species that are less

common than *T. loheri* and that have smaller population sizes or more restricted distributions. I therefore recommend the establishment of ecological corridors to increase genetic connectivity between the remaining forested areas with the aim of reducing the risk of inbreeding and loss of genetic diversity.

Nug-as forest in Alcoy and Boljoon, and the remaining fragmented forests in Argao and Dalaguete should also be awarded formal protective status to prevent further destruction, fragmentation, and degradation. However, protective status is not likely to be effective without community support. This is evident in the CCPL, which is formally a protected area, but is experiencing ongoing decline. In contrast, Nug-as forest is not a formally protected area and, as such, does not have a Protected Area Management Board. However, it is well managed by a local Peoples Organisation supported by an NGO (Philippines Biodiversity Conservation Foundation, Inc.) and has witnessed habitat regeneration in the recent decade.

This study resulted in the first information about patterns of genetic diversity and genetic connectivity among the few remaining forest areas in Cebu. As such, it provides hypotheses that could form the basis of future studies. For example, the findings of this study suggest that *T. loheri* might also show inbreeding and a lack of genetic connectivity among forested areas in Cebu that were not included in this study. To test these hypotheses and to further develop our understanding of the patterns of genetic diversity and connectivity among the remaining forest areas in Cebu with the aim of identifying Cebu-wide or area-specific conservation pressures that are shared among flora, conservation genetic studies of other species need to be carried out. Although it would be valuable to specifically target future conservation genetic research at rare and therefore potentially the most threatened species in Cebu's forests so that species-specific conservation action can be taken, such studies might not reveal habitat specific patterns, such as general patterns of genetic diversity and connectivity among Cebu's remaining forests. This is because the population sizes of rare species might be too small to obtain data from a sufficient number of individuals per population to reveal statistically well-supported patterns or their distribution might be too local to uncover island-wide patterns. Such data could, however, be obtained from common species that, like *T. loheri*, are confined to forest ecosystems or remnants thereof. Two primary candidates for this are *Dysoxylum pauciflorum* Merr. (Meliaceae) and *Garcinia rubra* Merr. (Clusiaceae). I encountered these species in sufficient numbers in all four areas that I included in my study. They are different life forms than *T. loheri* (a tree and a shrub, respectively). In addition, in contrast to *T. loheri* which is commonly present along forest edges as well as in its interior, *D. pauciflorum* and *G. rubra* were exclusively found in the

forest interior. These two species might therefore provide a different perspective of patterns of genetic diversity and connectivity among Cebu's forests.

REFERENCES

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *WIREs Computational Statistics 2*(4), 433-459. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/wics.101. doi:10.1002/wics.101

Abràmoff, M. D., Magalhães, P. J., & Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics international, 11*(7), 36-42.

Acosta, R., Natividad, M., & Razal, R. (2006). *One century of forest rehabilitation in the Philippines: Approaches, outcomes and lessons* (pp. 6-41, Rep.) (Pulhin J., Chokkalingam U., Peras R., Carandang A., Lasco R., & Toma T., Eds.). Center for International Forestry Research. Retrieved from www.jstor.org/stable/resrep02063.7

Adams, D. C., Rohlf, F. J., & Slice, D. E. (2004). Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology, 71*(1), 5-16.

Aguilar, R., Quesada, M., Ashworth, L., Herrerias‐Diego, Y., & Lobo, J. (2008). Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Molecular Ecology, 17*(24), 5177-5188.

Aizawa, M., Yoshimaru, H., Saito, H., Katsuki, T., Kawahara, T., Kitamura, K., Shi, F., Sabirov, R., & Kaji, M. (2009). Range-wide genetic structure in a north-east Asian spruce (*Picea jezoensis*) determined using nuclear microsatellite markers. *Journal of Biogeography, 36*(5), 996-1007. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2699.2008.02074.x. doi:10.1111/j.1365-2699.2008.02074.x

Alashwal, H., El Halaby, M., Crouse, J. J., Abdalla, A., & Moustafa, A. A. (2019). The Application of Unsupervised Clustering Methods to Alzheimer's Disease. *Frontiers In Computational Neuroscience 13*(31). Retrieved from https://www.frontiersin.org/article/10.3389/fncom.2019.00031. doi:10.3389/fncom.2019.00031

Arribas, P., Andújar, C., Sánchez‐Fernández, D., Abellán, P., & Millán, A. (2013). Integrative taxonomy and conservation of cryptic beetles in the Mediterranean region (Hydrophilidae). *Zoologica Scripta, 42*(2), 182-200.

Avise, J. C. (2004). *Molecular markers, natural history and evolution* (2nd ed.). USA: Sinauer Associates Inc.

Baack, E. J., & Rieseberg, L. H. (2007). A genomic view of introgression and hybrid speciation. *Current Opinion in Genetics & Development, 17*(6), 513-518.

Baldwin, B. G., & Markos, S. (1998). Phylogenetic Utility of the External Transcribed Spacer (ETS) of 18S–26S rDNA: Congruence of ETS and ITS Trees of *Calycadenia* (Compositae). *Molecular Phylogenetics and Evolution, 10*(3), 449-463.

Ballesteros-Mejia, L., Lima, N. E., Lima-Ribeiro, M. S., & Collevatti, R. G. (2016). Pollination mode and mating system explain patterns in genetic differentiation in neotropical plants. *PloS One, 11*(7), e0158660.

Ballian, D., Longauer, R., Mikić, T., Paule, L., Kajba, D., & Gömöry, D. (2006). Genetic structure of a rare European conifer, Serbian spruce (*Picea omorika* (Panč.) Purk.). *Plant Systematics and Evolution, 260*(1), 53-63. Retrieved from https://doi.org/10.1007/s00606-006-0431-z. doi:10.1007/s00606-006-0431-z

Bankoff, G. (2007). One island too many: reappraising the extent of deforestation in the Philippines prior to 1946. *Journal of Historical Geography, 33*(2), 314-334.

Barr, K. R., Kus, B. E., Preston, K. L., Howell, S., Perkins, E., & Vandergast, A. G. (2015). Habitat fragmentation in coastal southern California disrupts genetic connectivity in the cactus wren (*Campylorhynchus brunneicapillus*). *Molecular Ecology, 24*(10), 2349-2363. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.13176. doi:10.1111/mec.13176

Baum, D. A., & Shaw, K. L. (1995). Genealogical perspectives on the species problem. *Experimental and Molecular Approaches to Plant Biosystematics, 53*(289-303), 123-124.

Bena, G., Jubier, M.-F., Olivieri, I., & Lejeune, B. (1998). Ribosomal external and internal transcribed spacers: combined use in the phylogenetic analysis of *Medicago* (Leguminosae). *Journal of Molecular Evolution, 46*(3), 299-306.

Bensel, T. (2008). Fuelwood, deforestation, and land degradation: 10 years of evidence from Cebu province, the Philippines. *Land Degradation & Development, 19*(6), 587-605.

Bergl, R. A., & Vigilant, L. (2007). Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (*Gorilla gorilla diehli*). *Molecular Ecology, 16*(3), 501-516. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2006.03159.x. doi:10.1111/j.1365-294X.2006.03159.x

Blair, C., & Bryson Jr, R. W. (2017). Cryptic diversity and discordance in single‐locus species delimitation methods within horned lizards (Phrynosomatidae: Phrynosoma). *Molecular Ecology Resources, 17*(6), 1168-1182.

Blattner, F. R. (1999). Direct amplification of the entire ITS region from poorly preserved plant material using recombinant PCR. *BioTechniques, 27*(6), 1180-1186.

Blum, M. J., Bagley, M. J., Walters, D. M., Jackson, S. A., Daniel, F. B., Chaloud, D. J., & Cade, B. S. (2012). Genetic diversity and species diversity of stream fishes covary across a land-use gradient. *Oecologia, 168*(1), 83-95.

Bonhomme, V., Picq, S., Gaucherel, C., & Claude, J. (2014). Momocs: outline analysis using R. *Journal of Statistical Software, 56*(13), 1-24.

Bouckaert, R. (2019). Model selection with nested sampling: tutorial using BEAST v2.5.2. Retrieved from https://taming-the-beast.org/tutorials/NS-tutorial/NS-tutorial.pdf

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32.

Brizicky, G. K. (1965). The genera of Vitaceae in the southeastern United States. *Journal of the Arnold Arboretum, 46*(1), 48-67.

Brody, J. R., & Kern, S. E. (2004). Sodium boric acid: a Tris-free, cooler conductive medium for DNA electrophoresis. *Biotechniques*, *36*(2), 214-217.

Brookfield, J. F. Y. (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology, 5*(3), 453-455.

Cadiz, G. O., & Buot Jr, I. E. (2009). An Enumeration of the Woody Plants of Cantipla Forest Fragments, Cebu Island, Philippines. *Philippine Journal of Systematic Biology, 3,*1-7.

Calonje, M., Martín-Bravo, S., Dobeš, C., Gong, W., Jordon-Thaden, I., Kiefer, C., Kiefer, M., Paule, J., Schmickl, R., & Koch, M. A. (2009). Non-coding nuclear DNA markers in phylogenetic reconstruction. *Plant Systematics and Evolution, 282*(3), 257-280. Retrieved from https://doi.org/10.1007/s00606-008-0031-1. doi:10.1007/s00606-008-0031-1

Campoy, J. A., Lerigoleur-Balsemin, E., Christmann, H., Beauvieux, R., Girollet, N., Quero-García, J., Dirlewanger, E., & Barreneche, T. (2016). Genetic diversity, linkage disequilibrium, population structure and construction of a core collection of *Prunus avium* L. landraces and bred cultivars. *BMC plant biology, 16*(1), 49.

Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology, 22*(17), 4369-4383. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.12413. doi:10.1111/mec.12413

Chen, P., Chen, L., & Wen, J. (2011). The first phylogenetic analysis of *Tetrastigma* (Miq.) Planch., the host of Rafflesiaceae. *Taxon, 60*(2), 499-512.

Chitwood, D. H., & Otoni, W. C. (2017). Morphometric analysis of *Passiflora* leaves: the relationship between landmarks of the vasculature and elliptical Fourier descriptors of the blade. *GigaScience, 6*(1), giw008.

Chokkalingam, U., Carandang, A., Pulhin, J., Lasco, R., Peras, R., & Toma, T. (Eds.). (2006). *One century of forest rehabilitation in the Philippines: Approaches, outcomes and lessons* (pp. 1-5, Rep.). Center for International Forestry Research. Retrieved from www.jstor.org/stable/resrep02063.6.

Claridge, M. F., Dawah, H. A., & Wilson, M. R. (1997). *Species: the units of biodiversity*. Chapman and Hall Ltd.

Collevatti, R. G., Grattapaglia, D., & Hay, J. D. (2001). Population genetic structure of the endangered tropical tree species *Caryocar brasiliense*, based on variability at microsatellite loci. *Molecular Ecology, 10*(2), 349-356.

Cope, J. S., Corney, D., Clark, J. Y., Remagnino, P., & Wilkin, P. (2012). Plant species identification using digital morphometrics: A review. *Expert Systems with Applications, 39*(8), 7562-7573.

Cottontail, V. M., Kalko, E. K., Cottontail, I., Wellinghausen, N., Tschapka, M., Perkins, S. L., & Pinto, C. M. (2014). High local diversity of *Trypanosoma* in a common bat species, and implications for the biogeography and taxonomy of the *T. cruzi* clade. *PloS One, 9*(9), e108603.

Craft, K. J., & Ashley, M. V. (2007). Landscape genetic structure of bur oak (*Quercus macrocarpa*) savannas in Illinois. *Forest Ecology and Management, 239*(1), 13-20. Retrieved from http://www.sciencedirect.com/science/article/pii/S0378112706010899. doi:https://doi.org/10.1016/j.foreco.2006.11.005

Cronquist, A. (1978). *Once again, what is a species? Biosystematics in agriculture.* Paper presented at the Beltsville Symposia in Agricultural Research.

Dakin, E., & Avise, J. (2004). Microsatellite null alleles in parentage analysis. *Heredity, 93*(5), 504.

Dalstein, V., Eberle, J., Fabrizi, S., Etzbauer, C., & Ahrens, D. (2019). COI-based species delimitation in Indochinese *Tetraserica* chafers reveal hybridisation despite strong divergence in male copulation organs. *Organisms Diversity & Evolution, 19*(2), 277-286.

Damschen, E. I., Haddad, N. M., Orrock, J. L., Tewksbury, J. J., & Levey, D. J. (2006). Corridors increase plant species richness at large scales. *Science, 313*(5791), 1284-1286.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature methods, 9*(8), 772.

Dayanandan, S., Dole, J., Bawa, K., & Kesseli, R. (1999). Population structure delineated with microsatellite markers in fragmented populations of a tropical tree, *Carapa guianensis* (Meliaceae). *Molecular Ecology, 8*(10), 1585-1592.

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic biology, 56*(6), 879-886.

Debout, G. D. G., Doucet, J. L., & Hardy, O. J. (2011). Population history and gene dispersal inferred from spatial genetic structure of a Central African timber tree, *Distemonanthus benthamianus* (Caesalpinioideae). *Heredity, 106*(1), 88-99. Retrieved from https://doi.org/10.1038/hdy.2010.35. doi:10.1038/hdy.2010.35

Department of Environment and Natural Resources Administrative Order. (2017). Updated national list of threatened Philippine plants and their categories. *Department of Environment and Natural Resources Philippines*. Retrieved from https://denr.gov.ph/section-policies/viewrec.php?id=4957&main=1.

Després, L., Gielly, L., Redoutet, B., & Taberlet, P. (2003). Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability. *Molecular Phylogenetics and Evolution, 27*(2), 185-196.

Devos, A. K., Van Huffel, S., Simonetti, A. W., van der Graaf, M., Heerschap, A., & Buydens, L. M. (2007). Classification of Brain Tumours by Pattern Recognition of Magnetic Resonance Imaging and Spectroscopic Data. In *Outcome Prediction in Cancer* (pp. 285-318): Elsevier.

DeWoody, J., Nason, J. D., & Hipkins, V. D. (2006). Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes, 6*(4), 951-957.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology, 7*(1), 214.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution, 29*(8), 1969-1973.

Dubreuil, M., Riba, M., González-Martínez, S. C., Vendramin, G. G., Sebastiani, F., & Mayol, M. (2010). Genetic effects of chronic habitat fragmentation revisited: Strong genetic structure in a temperate tree, *Taxus baccata* (Taxaceae), with great dispersal capability. *American Journal of Botany, 97*(2), 303-310. Retrieved from https://bsapubs.onlinelibrary.wiley.com/doi/abs/10.3732/ajb.0900148. doi:10.3732/ajb.0900148

Duminil, J., & Di Michele, M. (2009). Plant species delimitation: a comparison of morphological and molecular markers. *Plant Biosystems, 143*(3), 528-542.

Duminil, J., Kenfack, D., Viscosi, V., Grumiau, L., & Hardy, O. J. (2012). Testing species delimitation in sympatric species complexes: the case of an African tropical tree, *Carapa* spp.(Meliaceae). *Molecular Phylogenetics and Evolution, 62*(1), 275-285.

Dutson, G. C., Magsalay, P. M., & Timmins, R. J. (1993). The rediscovery of the Cebu Flowerpecker *Dicaeum quadricolor*, with notes on other forest birds on Cebu, Philippines. *Bird Conservation International, 3*(3), 235-243.

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources, 4*(2), 359-361.

Elisens, W. J., & Nelson, A. D. (1993). Morphological and isozyme divergence in *Gambelia* (Scrophulariaceae): species delimitation and biogeographic relationships. *Systematic Botany*, 454-468.

Ellstrand, N. C., & Elam, D. R. (1993). Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics, 24*(1), 217-242.

Ereshefsky, M. (2007). Species, taxonomy, and systematics. In *Philosophy of biology* (pp. 403-427). Elsevier.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology, 14*(8), 2611-2620.

Everitt, B., & Wiley, I. (2011). *Cluster Analysis* (5th;5. Aufl.; ed.). Chichester, West Sussex: Wiley.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics, 131*(2), 479-491.

Ezard, T., Fujisawa, T., & Barraclough, T. (2009). splits: SPecies' LImits by Threshold Statistics. R package version 1.0-14/r31. In.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics, 164*(4), 1567-1587.

Falush, D., Stephens, M., & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology notes, 7*(4), 574-578.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution, 17*(6), 368-376.

Filippi, C. V., Aguirre, N., Rivas, J. G., Zubrzycki, J., Puebla, A., Cordes, D., Moreno, M. V., Fusari, C. M., Alvarez, D., & Heinz, R. A. (2015). Population structure and genetic diversity characterization of a sunflower association mapping population using SSR and SNP markers. *BMC Plant Biology, 15*(1), 52.

Fischer, J., & Lindenmayer, D. B. (2007). Landscape modification and habitat fragmentation: a synthesis. *16*(3), 265-280. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1466-8238.2007.00287.x. doi:10.1111/j.1466-8238.2007.00287.x

Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic biology, 20*(4), 406-416.

Fontaneto, D., Herniou, E. A., Boschetti, C., Caprioli, M., Melone, G., Ricci, C., & Barraclough, T. G. (2007). Independently evolving species in asexual bdelloid rotifers. *PLoS biology, 5*(4), e87.

Fontaneto, D., Iakovenko, N., Eyres, I., Kaya, M., Wyman, M., & Barraclough, T. G. (2011). Cryptic diversity in the genus *Adineta* Hudson & Gosse, 1886 (Rotifera: Bdelloidea: Adinetidae): a DNA taxonomy approach. *Hydrobiologia, 662*(1), 27-33.

Food and Agriculture Organization. (1997). Asia-Pacific Forestry Sector Outlook Study: Commentary on Forest Policy in the Asia-Pacific Region (A Review for Indonesia, Malaysia, New Zealand, Papua New Guinea, Philippines, Thailand and Western Samoa). Retrieved from http://www.fao.org/3/w7730e/w7730e00.htm#Contents

Forest Management Bureau. (2019). Philippine forests at a glance. Retrieved from https://drive.google.com/file/d/1poFM4C-2BJSzFlSn6oN3io7xQZkjg2Nj/view?usp=embed_facebook

Frades, I., & Matthiesen, R. (2010). Overview on techniques in cluster analysis. In *Bioinformatics methods in clinical research* (pp. 81-107): Springer.

Fraley, C., & Raftery, A. E. (1999). MCLUST: Software for model-based cluster analysis. *Journal of classification, 16*(2), 297-306.

Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: a review. *Genetics Research, 66*(2), 95-107.

Frankham, R., Briscoe, D. A., & Ballou, J. D. (2010). *Introduction to conservation genetics* (2nd ed.): Cambridge university press.

Fujisawa, T., & Barraclough, T. G. (2013). Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. *Systematic biology, 62*(5), 707-724. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3739884/pdf/syt033.pdf.

Fuller, D. Q., & Hickey, L. J. (2005). Systematics and leaf architecture of the Gunneraceae. *The Botanical Review, 71*(3), 295.

Gadiana, L. P. M. P. M. (2004). The Philippine bare-backed fruit bat *Dobsonia chapmani* Rabor, 1952: Rediscovery and conservation status on Cebu Island. *Silliman Journal, 45*(2).

Gagnepain, F. (1910). *Tetrastigma* (Ampélidacées) nouveaux ou peu connus. *Notulae Systematicae, 1*: 261–271.

Gan, G., Ma, C. P. D., & Wu, J. (2007). *Data clustering: theory, algorithms, and applications* (Vol. 20). Alexandria, Va;Philadelphia, Pa;: SIAM, Society for Industrial and Applied Mathematics.

García-Melo, J. E., Oliveira, C., Silva, G. J. D. C., Ochoa-Orrego, L. E., Pereira, L. H. G., & Maldonado-Ocampo, J. A. (2019). Species delimitation of neotropical *Characins* (Stevardiinae): Implications for taxonomy of complex groups. *PloS One, 14*(6), e0216786.

Garnier‐Géré, P., & Chikhi, L. (2001). Population subdivision, Hardy–Weinberg equilibrium and the Wahlund effect. In eLS, John Wiley & Sons, Ltd (Ed.). doi:10.1002/9780470015902.a0005446.pub3.

Global Forest Watch. (2019). Intact forest in the Philippines. Retrieved from www.globalforestwatch.org

Goncalves, A. L., García, M. V., Heuertz, M., & González-Martínez, S. C. (2019). Demographic history and spatial genetic structure in a remnant population of the subtropical tree *Anadenanthera colubrina* var. *cebil* (Griseb.) Altschul (Fabaceae). *Annals of Forest Science, 76*(1), 18. Retrieved from https://doi.org/10.1007/s13595-019-0797-z. doi:10.1007/s13595-019-0797-z

González-Martínez, S. C., Dubreuil, M., Riba, M., Vendramin, G. G., Sebastiani, F., & Mayol, M. (2010). Spatial genetic structure of *Taxus baccata* L. in the western Mediterranean Basin: Past and present limits to gene movement over a broad geographic scale. *Molecular Phylogenetics and Evolution, 55*(3), 805-815. Retrieved from http://www.sciencedirect.com/science/article/pii/S1055790310000990. doi:https://doi.org/10.1016/j.ympev.2010.03.001

Gonzalez, J., Dans, A., Pedregosa, M., Chiu, C., & Villahermosa, R. (1999). Island-wide survey of forest and fauna and flora inventory of selected sites for priority conservation on Cebu. *Manila: Fauna & Flora International (unpublished)*.

Goverde, M., Schweizer, K., Baur, B., & Erhardt, A. (2002). Small-scale habitat fragmentation effects on pollinator behaviour: experimental evidence from the bumblebee *Bombus veteranus* on calcareous grasslands. *Biological Conservation, 104*(3), 293-299.

Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 857-871.

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic biology, 52*(5), 696-704.

Habib, S., Dang, V.-C., Ickert-Bond, S. M., Zhang, J.-L., Lu, L.-M., Wen, J., & Chen, Z.-D. (2017). Robust phylogeny of *Tetrastigma* (Vitaceae) based on ten plastid DNA regions: Implications for infrageneric classification and seed character evolution. *Frontiers in Plant Science, 8*, 590. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5405133/pdf/fpls-08-00590.pdf.

Hale, M. L., Burg, T. M., & Steeves, T. E. J. P. o. (2012). Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. *7*(9), e45170.

Hall, L. A., & Beissinger, S. R. (2014). A practical toolbox for design and analysis of landscape genetics studies. *Landscape Ecology, 29*(9), 1487-1504.

Hamilton, A. J. (2005). Species diversity or biodiversity? *Journal of Environmental Management, 75*(1), 89-92. Retrieved from http://www.sciencedirect.com/science/article/pii/S0301479705000149. doi:https://doi.org/10.1016/j.jenvman.2004.11.012

Hamilton, M. (1999). Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular ecology, 8*, 521-523.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science, 342*(6160), 850. Retrieved from http://science.sciencemag.org/content/342/6160/850.abstract. doi:10.1126/science.1244693

Harris, R. J., & Reed, J. M. (2002). *Behavioral barriers to non-migratory movements of birds.* Paper presented at the Annales Zoologici Fennici.

Harrison, R. G., & Larson, E. L. (2014). Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity, 105*(S1), 795-809.

Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution, 59*(8), 1633-1638.

Hickey, L. J., & Taylor, D. W. (1991). The leaf architecture of *Ticodendron* and the application of foliar characters in discerning its relationships. *Annals of the Missouri Botanical Garden*, 105-130.

Hickey, L. J., & Wolfe, J. A. (1975). The bases of angiosperm phylogeny: vegetative morphology. *Annals of the Missouri Botanical Garden*, 538-589.

Holderegger, R., Kamm, U., & Gugerli, F. (2006). Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology, 21*(6), 797-807. Retrieved from https://doi.org/10.1007/s10980-005-5245-9. doi:10.1007/s10980-005-5245-9

Honnay, O., & Bossuyt, B. (2005). Prolonged clonal growth: escape route or route to extinction? *Oikos, 108*(2), 427-432.

Huelsenbeck, J. P., Ronquist, F., Nielsen, R., & Bollback, J. P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. *Science, 294*(5550), 2310-2314.

Hufford, K. M., Mazer, S. J. J. T. i. E., & Evolution. (2003). Plant ecotypes: genetic differentiation in the age of ecological restoration. *18*(3), 147-155.

Jakosalem, P. G. C., Collar, N. J., & Gill, J. A. (2013). Habitat selection and conservation status of the endemic Ninox hawk-owl on Cebu, Philippines. *Bird Conservation International, 23*(3), 360-370.

Jamieson, I. G., Grueber, C. E., Waters, J. M., & Gleeson, D. M. (2008). Managing genetic diversity in threatened populations: a New Zealand perspective. *New Zealand Journal of Ecology*, 130-137.

Janson, C. H. (1983). Adaptation of fruit morphology to dispersal agents in a neotropical forest. *Science, 219*(4581), 187-189.

Jenkins, T. L., & Stevens, J. R. (2018). Assessing connectivity between MPAs: Selecting taxa and translating genetic data to inform policy. *Marine Policy, 94*, 165-173.

Jensen, R. J. (2003). The conundrum of morphometrics. *Taxon, 52*(4), 663-671.

Jensen, R. J., Ciofani, K. M., & Miramontes, L. C. (2002). Lines, outlines, and landmarks: morphometric analyses of leaves of *Acer rubrum, Acer saccharinum* (Aceraceae) and their hybrid. *Taxon, 51*(3), 475-492.

Joly, S., & Bruneau, A. (2007). Delimiting species boundaries in *Rosa* sect. *Cinnamomeae* (Rosaceae) in eastern North America. *Systematic Botany*, *32*(4), 819-836.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics, 24*(11), 1403-1405.

Jombart, T., & Collins, C. (2015). A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0. 0. *London: Imperial College London, MRC Centre for Outbreak Analysis and Modelling*.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics, 11*(1), 94.

Kahilainen, A., Puurtinen, M., & Kotiaho, J. S. (2014). Conservation implications of species–genetic diversity correlations. *Global Ecology and Conservation, 2*, 315-323.

Kalinowski, S. T. (2005). hp‐rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular ecology notes, 5*(1), 187-189.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773-795.

Kassambara, A., & Mundt, F. (2016). Factoextra: extract and visualize the results of multivariate data analyses. *R package version, 1*(3), 2016.

Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis* (Vol. 344): John Wiley & Sons.

Klingenberg, C., Duttke, S., Whelan, S., & Kim, M. (2012). Developmental plasticity, morphological variation and evolvability: a multilevel analysis of morphometric integration in the shape of compound leaves. *Journal of Evolutionary Biology, 25*(1), 115-129.

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources, 11*(2), 353-357.

Koressaar, T., & Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics, 23*(10), 1289-1291.

Kramer, A. T., Ison, J. L., Ashley, M. V., & Howe, H. F. (2008). The paradox of forest fragmentation genetics. *Conservation Biology, 22*(4), 878-885.

Kummer, D., Concepcion, R., & Canizares, B. (1994). Environmental degradation in the uplands of Cebu. *Geographical Review*, 266-276.

Lang, A. S., Bocksberger, G., & Stech, M. (2015). Phylogeny and species delimitations in European *Dicranum* (Dicranaceae, Bryophyta) inferred from nuclear and plastid DNA. *Molecular Phylogenetics and Evolution, 92*, 217-225.

Larson, E. R., Castelin, M., Williams, B. W., Olden, J. D., & Abbott, C. L. (2016). Phylogenetic species delimitation for crayfishes of the genus *Pacifastacus*. *PeerJ. 4*, e1915.

Lasco, R. D., Visco, R. G., & Pulhin, J. M. (2001). Secondary forests in the Philippines: formation and transformation in the 20th century. *Journal of Tropical Forest Science*, 652-670.

Latiff, A. (2001). Studies in Malesian Vitaceae XII: Taxonomic notes on *Cissus, Ampelocissus, Nothocissus* and *Tetrastigma* and other genera. *Folia Malaysiana, 2*(3), 179-189.

Laurance, W. F. (2007). Forest destruction in tropical Asia. *Current Science*, 1544-1550.

Le Ru, B. P., Capdevielle-Dulac, C., Toussaint, E. F., Conlong, D., Van den Berg, J., Pallangyo, B., Ong'amo, G., Chipabika, G., Molo, R., & Overholt, W. A. (2014). Integrative taxonomy of *Acrapex* stem borers (Lepidoptera: Noctuidae: Apameini): combining morphology and Poisson Tree Process analyses. *Invertebrate Systematics, 28*(5), 451-475.

Leaf Architecture Working Group. (1999). Manual of leaf architecture. *Smithsonian Institution*.

Lemes, M. R., Gribel, R., Proctor, J., & Grattapaglia, D. (2003). Population genetic structure of mahogany (*Swietenia macrophylla* King, Meliaceae) across the Brazilian Amazon, based on variation at microsatellite loci: implications for conservation. *Molecular Ecology, 12*(11), 2875-2883.

Levey, D. J., Bolker, B. M., Tewksbury, J. J., Sargent, S., & Haddad, N. M. (2005). Effects of landscape corridors on seed dispersal by birds. *Science, 309*(5731), 146-148.

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic biology, 50*(6), 913-925.

Li, L., & Yan, H. (2013). A remarkable new species of *Liparis* (Orchidaceae) from China and its phylogenetic implications. *PloS One, 8*(11), e78112.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news, 2*(3), 18-22.

Lienert, J. (2004). Habitat fragmentation effects on fitness of plant populations–a review. *Journal for Nature Conservation, 12*(1), 53-72.

Lillo, E. P., Alcazar, S. M. T., Nuevo, R. U., & Malaki, A. B. B. (2015). Vascular Plants of Mangrove Forest in Argao, Cebu, Philippines. *Tropical Technology Journal, 18*(1), 1-9.

Lillo, E. P., & Buot Jr, I. E. Species composition of Argao mangrove forest, Cebu, Philippines. *Journal of Wetlands Biodiversity, 6,* 37-45.

Linder, C. R., Goertzen, L. R., Heuvel, B. V., Francisco-Ortega, J., & Jansen, R. K. (2000). The complete external transcribed spacer of 18S-26S rDNA: amplification and phylogenetic utility at low taxonomic levels in Asteraceae and closely allied families. *Molecular Phylogenetics and Evolution, 14*(2), 285-303.

Linhart, Y. B., & Grant, M. C. (1996). Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics, 27*(1), 237-277.

Logan, S. A., Phuekvilai, P., & Wolff, K. (2015). Ancient woodlands in the limelight: delineation and genetic structure of ancient woodland species *Tilia cordata* and *Tilia platyphyllos* (Tiliaceae) in the UK. *Tree Genetics & Genomes, 11*(3), 52. Retrieved from https://doi.org/10.1007/s11295-015-0872-z. doi:10.1007/s11295-015-0872-z

Lowe, A., Boshier, D., Ward, M., Bacles, C., & Navarro, C. (2005). Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for neotropical trees. *Heredity, 95*(4), 255.

Lowe, W. H., & Allendorf, F. W. (2010). What can genetics tell us about population connectivity? *Molecular Ecology, 19*(15), 3038-3051. Retrieved from

https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2010.04688.x. doi:10.1111/j.1365-294X.2010.04688.x

Luo, A., Ling, C., Ho, S. Y., & Zhu, C.-D. (2018). Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology, 67*(5), 830-846. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6101526/pdf/syy011.pdf.

Ma, W.-J., Santos del Blanco, L., & Pannell, J. R. (2019). A new biological species in the *Mercurialis annua* polyploid complex: functional divergence in inflorescence morphology and hybrid sterility. *Annals of Botany, 124*(1), 165-178. Retrieved from https://doi.org/10.1093/aob/mcz058. doi:10.1093/aob/mcz058

MacQueen, J. (1967). *Some methods for classification and analysis of multivariate observations.* Paper presented at the Proceedings of the fifth Berkeley symposium on mathematical statistics and probability.

Magsalay, P. (1993). Rediscovery of four Cebu endemic birds (Philippines). *Asian International Journal of Life Sciences 2*(2)*,* 141-148.

Malaki, A. B. B. (2016). Population density assessment of Black Shama (*Copsychus cebuensis* Steere) in selected forest patches within the key biodiversity areas in Argao, Cebu Island, Philippines. *Asia Life Sciences, 25*(1), 325-335.

Malaki, A. B. B., Cruz, R. V. O., Bantayan, N. C., Racelis, D. A., Buot, I. E., & Florece, L. M. (2013). Landscape Pattern Impacts on the Population Density and Distribution of Black Shama (*Copsychus cebuensis* Steere) in Argao Watershed Reserve, Argao, Cebu, Philippines. *ISRN Biodiversity, 2013* Article ID 568498, 10 pages, 2013. Retrieved fromhttps://doi.org/10.1155/2013/568498.

Manen, J.-F., Natali, A., & Ehrendorfer, F. (1994). Phylogeny of Rubiaceae-Rubieae inferred from the sequence of a cpDNA intergene region. *Plant Systematics and Evolution, 190*(3-4), 195-211.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer research, 27*(2 Part 1), 209-220.

Manthey, J. D., & Moyle, R. G. (2015). Isolation by environment in White-breasted Nuthatches (*Sitta carolinensis*) of the Madrean Archipelago sky islands: A landscape genomics approach. *Molecular Ecology*, *24*(14), 3628-3638. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.13258. doi:10.1111/mec.13258

Markos, S., & Baldwin, B. G. (2001). Higher-level relationships and major lineages of *Lessingia* (Compositae, Astereae) based on nuclear rDNA internal and external transcribed spacer (ITS and ETS) sequences. *Systematic Botany, 26*(1), 168-184.

Matos, E., Oliveira, E., Jesus, O., & Dantas, J. (2013). Microsatellite markers of genetic diversity and population structure of *Carica papaya*. *Annals of Applied Biology, 163*(2), 298-310.

Mayr, E. (1942). *Systematics and the Origin of Species*. Columbia University Press: New York.

Meirmans, P. G. (2012). The trouble with isolation by distance. *Molecular Ecology, 21*(12), 2839-2846. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2012.05578.x. doi:10.1111/j.1365-294X.2012.05578.x

Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: FST and related measures. *Molecular ecology resources, 11*(1), 5-18.

Mello, B., Vilela, J. F., & Schrago, C. G. (2018). Conservation phylogenetics and computational species delimitation of Neotropical primates. *Biological conservation, 217*, 397-406.

Merrill, E. D. (1912). Sertulum Bontocense: New Or Interesting Plants Collected in Bontoc Subprovince, Luzon by Father Morice Vanoverbergh. *Philippine Journal of Science, C 7*, 71-107.

Merrill, E. D. (1914). Plantae Wenzelianae. *Philippine Journal of Science, C 9,* 370-371.

Merrill, E. D. (1916). New or interesting Philippine Vitaceae. *Philippine Journal of Science, C 11*, 125-145.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). *Creating the CIPRES Science Gateway for inference of large phylogenetic trees.* Paper presented at the 2010 gateway computing environments workshop (GCE).

Miquel, F. A. W. (1861). Sumatra, Zijne Plantenwereld en hare voortbrengselen. *Flora van Nederlandsch Indië, 3*, 515.

Modica, M. V., Puillandre, N., Castelin, M., Zhang, Y., & Holford, M. (2014). A good compromise: rapid and robust species proxies for inventorying biodiversity hotspots using the Terebridae (Gastropoda: Conoidea). *PloS one, 9*(7), e102160. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4086986/pdf/pone.0102160.pdf.

Moffat, C. E., Ensing, D. J., Gaskin, J. F., De Clerck-Floate, R. A., & Pither, J. (2015). Morphology delimits more species than molecular genetic clusters of invasive *Pilosella*. *American Journal of Botany, 102*(7), 1145-1159. Retrieved from https://bsapubs.onlinelibrary.wiley.com/doi/full/10.3732/ajb.1400466.

Monaghan, M. T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D. J., Lees, D. C., Ranaivosolo, R., Eggleton, P., & Barraclough, T. G. (2009). Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic biology, 58*(3), 298-311.

Múrria, C., Bonada, N., Arnedo, M. A., Prat, N., & Vogler, A. P. (2013). Higher β-and γ-diversity at species and genetic levels in headwaters than in mid-order streams in *Hydropsyche* (Trichoptera). *Freshwater Biology, 58*(11), 2226-2236.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature, 403*(6772), 853.

Narum, S. R. (2006). Beyond Bonferroni: less conservative analyses for conservation genetics. *ConservationGenetics, 7*(5), 783-787.

Nery, E. K., & Fiaschi, P. (2019). Geometric Morphometrics Dismiss the Polymorphic *Hydrocotyle quinqueloba* (Araliaceae) from the Neotropics. *Systematic Botany, 44*(2), 451-469.

Nickrent, D. L., Costea, M., Barcelona, J. F., Pelser, P. B., & Nixon, K. (2006 onwards). PhytoImages. Retrieved from Available from: http://www.phytoimages.siu.edu.

Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E., Mathesius, U., Poot, P., Purugganan, M. D., Richards, C. L., & Valladares, F. (2010). Plant phenotypic plasticity in a changing climate. *Trends in Plant Science, 15*(12), 684-692.

Nikolić, T. (1995). Numerical taxonomy of the Ranunculaceae. In *Systematics and Evolution of the Ranunculiflorae* (pp. 281-282): Springer.

Nybom, H. (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology, 13*(5), 1143-1155.

Obertegger, U., Fontaneto, D., & Flaim, G. (2012). Using DNA taxonomy to investigate the ecological determinants of plankton diversity: explaining the occurrence of *Synchaeta* spp.(Rotifera, Monogononta) in mountain lakes. *Freshwater Biology, 57*(8), 1545-1553.

Okuyama, Y., Fujii, N., Wakabayashi, M., Kawakita, A., Ito, M., Watanabe, M., Murakami, N., & Kato, M. (2004). Nonuniform concerted evolution and chloroplast capture:

heterogeneity of observed introgression patterns in three molecular data partition phylogenies of Asian *Mitella* (Saxifragaceae). *Molecular Biology and Evolution, 22*(2), 285-296.

Ouinsavi, C., Sokpon, N., & Khasa, D. P. (2009). Genetic diversity and population structure of a threatened African tree species, *Milicia excelsa*, using nuclear microsatellites DNA markers. *International Journal of Forestry Research, 2009*. Article ID 210179.

Oxelman, B., Lidén, M., & Berglund, D. (1997). Chloroplastrps16 intron phylogeny of the tribe *Sileneae* (Caryophyllaceae). *Plant Systematics and Evolution, 206*(1-4), 393-410.

Paguntalan, L. J., Lepitan-Tabao, M., Jakosalem, P. G. C., Orolfo, A. O., & Rica, A. N. (2015). *State of Cebu's Biodiversity*. Philippines Biodiversity Conservation Foundation, Inc. Bacolod City, Philippines.

Paguntalan, L. M. J., & Jakosalem, P. G. (2008a). Significant records of birds in forests on Cebu island, central Philippines. *Forktail, 24*, 48-56.

Paguntalan, L. M. J., & Jakosalem, P. G. C. (2008b). Conserving threatened and endemic fruit bats in isolated forest patches in Cebu (with notes on new records and rediscoveries). *Silliman Journal, 48*(2), 81-94.

Papadopoulou, A., Anastasiou, I., Keskin, B., & Vogler, A. P. (2009). Comparative phylogeography of tenebrionid beetles in the Aegean archipelago: the effect of dispersal ability and habitat preference. *Molecular Ecology, 18*(11), 2503-2517. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2009.04207.x. doi:10.1111/j.1365-294X.2009.04207.x

Papadopoulou, A., Bergsten, J., Fujisawa, T., Monaghan, M. T., Barraclough, T. G., & Vogler, A. P. (2008). Speciation and DNA barcodes: testing the effects of dispersal on the formation of discrete sequence clusters. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1506), 2987-2996.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics (Oxford, England), 28*(19), 2537-2539. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/22820204

Pelser, P., Barcelona, J., & Nickrent, D (eds.). 2011 onwards. Co's Digital Flora of the Philippines. http://www.philippineplants.org.

Pelser, P., Nickrent, D., Barcelona, J. (2018). A conservation genetic study of *Rafflesia speciosa* (Rafflesiaceae): patterns of genetic diversity and differentiation within and

between islands. *Blumea-Biodiversity, Evolution and Biogeography of Plants, 63*(2), 93-101.

Pelser, P. B., Doble, K. J. S., O'Byrne, P., Ormerod, P., & Barcelona, J. F. (2016a). *Gastrodia cajanoae* (Orchidaceae: Epidendroideae: Gastrodieae), a new species from the Philippines. *Phytotaxa, 266*(1), 53-56.

Pelser, P. B., Nickrent, D. L., & Barcelona, J. F. (2016b). Untangling a vine and its parasite: Host specificity of Philippine *Rafflesia* (Rafflesiaceae). *Taxon, 65*(4), 739-758.

Pelser, P. B., Nickrent, D. L., Gemmill, C. E., & Barcelona, J. F. (2017). Genetic diversity and structure in the Philippine *Rafflesia lagascae* complex (Rafflesiaceae) inform its taxonomic delimitation and conservation. *Systematic botany*, *42*(3), 543-553.

Pelser, P. B., Nickrent, D. L., Reintar, A. R., & Barcelona, J. F. (2016c). *Lepeostegeres cebuensis* (Loranthaceae), a new mistletoe species from Cebu, Philippines. *Phytotaxa, 266*(1), 048-052.

Petrini, O., & Fisher, P. J. (1988). A comparative study of fungal endophytes in xylem and whole stem of *Pinus sylvestris* and *Fagus sylvatica*. *Transactions of the British Mycological Society, 91*(2), 233-238. Retrieved from http://www.sciencedirect.com/science/article/pii/S0007153688802109. doi:https://doi.org/10.1016/S0007-1536(88)80210-9

Poczai, P., & Hyvönen, J. (2010). Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. *Molecular Biology Reports, 37*(4), 1897-1912.

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D., & Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology, 55*(4), 595-609.

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, A., & Lareu, M. V. (2013). An overview of STRUCTURE: applications, parameter settings, and supporting software. *Frontiers in Genetics, 4*, 98-98. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23755071

Potapov, P., Hansen, M. C., Laestadius, L., Turubanova, S., Yaroshenko, A., Thies, C., Smith, W., Zhuravleva, I., Komarova, A., Minnemeyer, S., & Esipova, E. (2017). The last frontiers of wilderness: Tracking loss of intact forest landscapes from 2000 to 2013. *Science Advances, 3*(1), e1600821-e1600821. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5235335/. doi:10.1126/sciadv.1600821

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945-959.

Rabor, D. (1959). The impact of deforestation on birds of Cebu, Philippines, with new records for that island. *The Auk, 76*(1), 37-43.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology, 67*(5), 901-904.

Rathcke, B. J., & Jules, E. S. J. C. S. (1993). Habitat fragmentation and plant–pollinator interactions. *Current Science 65(3),* 273-277.

Raymond, M., & Rousset, F. (1995). An exact test for population differentiation. *Evolution, 49*(6), 1280-1283.

Reddy, P. A., Gour, D. S., Bhavanishankar, M., Jaggi, K., Hussain, S. M., Harika, K., & Shivaji, S. (2012). Genetic Evidence of Tiger Population Structure and Migration within an Isolated and Fragmented Landscape in Northwest India. *PloS One, 7*(1), e29827. Retrieved from https://doi.org/10.1371/journal.pone.0029827. doi:10.1371/journal.pone.0029827

Ricketts, T. H. (2001). The matrix matters: effective isolation in fragmented landscapes. *The American Naturalist, 158*(1), 87-99.

Rieseberg, L. H. (1991). Hybridization in rare plants: insights from case studies in *Cercocarpus* and *Helianthus*. *Genetics and Conservation of Rare Plants*, 171-181.

Rohlf, F. J., & Marcus, L. F. (1993). A revolution morphometrics. *Trends in Ecology & Evolution, 8*(4), 129-132.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology, 61*(3), 539-542.

Rossetto, M., Jackes, B. R., Scott, K. D., & Henry, R. J. (2002). Is the genus *Cissus* (Vitaceae) monophyletic? Evidence from plastid and nuclear ribosomal DNA. *Systematic Botany, 27*(3), 522-534.

Rossetto, M., Slade, R., Baverstock, P. R., Henry, R. J., & Lee, L. S. (1999). Microsatellite variation and assessment of genetic structure in tea tree (*Melaleuca alternifolia*– Myrtaceae). *Molecular Ecology, 8*(4), 633-643.

RStudio Team. (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA, USA. In.

Russel, P. M., Brewer, B. J., Klaere, S., & Bouckaert, R. R. (2018). Model selection and parameter inference in phylogenetics using Nested Sampling. *Systematic Biology, 68*(2), 219-233.

Saar, D. E., Polans, N. O., & Sørensen, P. D. (2003). A phylogenetic analysis of the genus *Dahlia* (Asteraceae) based on internal and external transcribed spacer regions of nuclear ribosomal DNA. *Systematic Botany, 28*(3), 627-640.

Safner, T., Miller, M. P., McRae, B. H., Fortin, M.-J., & Manel, S. (2011). Comparison of Bayesian clustering and edge detection methods for inferring boundaries in landscape genetics. *International Journal of Molecular Sciences, 12*(2), 865-889.

Salares, V. B., Obico, J. J. A., Ormerod, P., Barcelona, J. F., & Pelser, P. B. (2018). Taxonomic novelties from Cebu: a new species of *Vaccinium* (Ericaceae) and a new record of *Phaius* (Orchidaceae) for the Philippines. *Phytotaxa, 360*(3), 255-262.

Sallares, R., & Brown, T. A. (2004). Phylogenetic analysis of complete 5′ external transcribed spacers of the 18S ribosomal RNA genes of diploid *Aegilops* and related species (Triticeae, Poaceae). *Genetic Resources and Crop Evolution, 51*(7), 701-712.

Seidenschwarz, F. (1988). Forest types of Cebu island. *Philippine Quarterly of Culture and Society, 16*(2), 93-105. Retrieved from http://www.jstor.org/stable/29791948.

Seidenschwarz, F. (2013). *Cynometra cebuensis*, a new species of Leguminosae (Caesalpinioidae) from the Philippines. *Blumea-Biodiversity, Evolution and Biogeography of Plants, 58*(1), 18-20.

Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology letters, 9*(5), 615-629.

Shipunov, A., Kosenko, Y., & Volkova, P. (2011). Floral polymorphism in common primrose (*Primula vulgaris* Huds., Primulaceae) of the Northeastern Black Sea coast. *Plant Systematics and Evolution, 296*(3-4), 167-178.

Simmons, M. P., & Ochoterena, H. (2000). Gaps as characters in sequence-based phylogenetic analyses. *Systematic biology*, *49*(2), 369-381.

Sinnock, P. (1975). The Wahlund effect for the two-locus model. *The American Naturalist*, *109*(969), 565-570.

Skoracka, A., Kuczyński, L., Rector, B., & Amrine Jr, J. W. (2014). Wheat curl mite and dry bulb mite: untangling a taxonomic conundrum through a multidisciplinary approach. *Biological Journal of the Linnean Society, 111*(2), 421-436.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science, 236*(4803), 787. Retrieved from http://science.sciencemag.org/content/236/4803/787.abstract. doi:10.1126/science.3576198

Song, C., Wang, Q., Zhang, R., Sun, B., & Wang, X. (2016). Exploring the utility of DNA barcoding in species delimitation of *Polypedilum* (Tripodura) non-biting midges (Diptera: Chironomidae). *Zootaxa, 4079*(5), 534-550.

Sork, V. L., & Smouse, P. E. (2006). Genetic analysis of landscape connectivity in tree populations. *Landscape Ecology, 21*(6), 821-836. Retrieved from https://doi.org/10.1007/s10980-005-5415-9. doi:10.1007/s10980-005-5415-9

Stahl, D., & Sallis, H. (2012). Model‐based cluster analysis. *Wiley Interdisciplinary Reviews: Computational Statistics, 4*(4), 341-358.

Stojnić, S., V Avramidou, E., Fussi, B., Westergren, M., Orlović, S., Matović, B., Trudić, B., Kraigher, H., A Aravanopoulos, F., & Konnert, M. (2019). Assessment of genetic diversity and population genetic structure of Norway Spruce (*Picea abies* (L.) Karsten) at Its Southern Lineage in Europe: implications for conservation of forest genetic resources. *Forests, 10*(3), 258.

Struebig, M. J., Kingston, T., Petit, E. J., Le Comber, S. C., Zubaid, A., Mohd‐Adnan, A., & Rossiter, S. J. (2011). Parallel declines in species and genetic diversity in tropical forest fragments. *Ecology Letters, 14*(6), 582-590.

Suchard, M. A., & Rambaut, A. (2009). Many-core algorithms for statistical phylogenetics. *Bioinformatics, 25*(11), 1370-1376.

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences, 114*(7), 1607-1612.

Taberlet, P., Gielly, L., Pautou, G., & Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant molecular biology, 17*(5), 1105-1109.

Taberlet, P., Zimmermann, N. E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., Niklfeld, H., Coldea, G., Mirek, Z., & Moilanen, A. (2012). Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecology Letters, 15*(12), 1439-1448.

Talavera, G., Dincă, V., & Vila, R. (2013). Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods in Ecology and Evolution, 4*(12), 1101-1110.

Tamaki, I., Kawashima, N., Setsuko, S., Lee, J.-H., Itaya, A., Yukitoshi, K., & Tomaru, N. (2019). Population genetic structure and demography of *Magnolia kobus*: variety borealis is not supported genetically. *Journal of Plant Research, 132*(6), 741-758. Retrieved from https://doi.org/10.1007/s10265-019-01134-6. doi:10.1007/s10265-019-01134-6

Tang, C. Q., Humphreys, A. M., Fontaneto, D., & Barraclough, T. G. (2014). Effects of phylogenetic reconstruction method on the robustness of species delimitation using single-locus data. *Methods in Ecology and Evolution, 5*(10), 1086-1094. Retrieved from http://spiral.imperial.ac.uk/bitstream/10044/1/18180/2/Methods%20in%20Ecology%20and%20Evolution_5_10_2014.pdf.

Templeton, A. R., Shaw, K., Routman, E., & Davis, S. K. (1990). The genetic consequences of habitat fragmentation. *Annals of the Missouri Botanical Garden 1*, 13-27.

Tewksbury, J. J., Levey, D. J., Haddad, N. M., Sargent, S., Orrock, J. L., Weldon, A., Danielson, B. J., Brinkerhoff, J., Damschen, E. I., & Townsend, P. (2002). Corridors affect plants, animals, and their interactions in fragmented landscapes. *Proceedings of the National Academy of Sciences, 99*(20), 12923-12926.

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63*(2), 411-423.

Tjaden, B., & Cohen, J. (2006). A survey of computational methods used in microarray data interpretation. In *Applied Mycology and Biotechnology* (Vol. 6, pp. 161-178): Elsevier.

Toussaint, E. F., Morinière, J., Müller, C. J., Kunte, K., Turlin, B., Hausmann, A., & Balke, M. (2015). Comparative molecular species delimitation in the charismatic Nawab butterflies (Nymphalidae, Charaxinae, Polyura). *Molecular Phylogenetics and Evolution, 91*, 194-209.

Turchetto, C., Mäder, G., Cazé, A. L. R., & Freitas, L. B. (2018). Genetic variability and population structure of *Passiflora contracta*, a bat-pollinated species from a fragmented rainforest. *Botanical Journal of the Linnean Society, 186*(2), 247-258.

Turland, N. J., Wiersema, J. H., Barrie, F. R., Greuter, W., Hawksworth, D., Herendeen, P. S., Knapp, S., Kusber, W.-H., Li, D.-Z., & Marhold, K. (2018). *International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017*: Koeltz Botanical Books.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic acids research*, *40*(15), e115-e115.

Uriarte, M., Anciães, M., da Silva, M. T. B., Rubim, P., Johnson, E., & Bruna, E. M. (2011). Disentangling the drivers of reduced long-distance seed dispersal by birds in an experimentally fragmented landscape. *Ecology, 92*(4), 924-937. Retrieved from https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/10-0709.1. doi:10.1890/10-0709.1

Vaishnav, V., Mahesh, S., & Kumar, P. (2019). Assessment of genetic structure of the endangered forest species *Boswellia serrata* population in Central India. *Journal of Tropical Forest Science, 31*(2), 200-210. Retrieved from www.jstor.org/stable/26626991.

Viscosi, V., & Cardini, A. (2011). Leaf morphology, taxonomy and geometric morphometrics: a simplified protocol for beginners. *PloS One, 6*(10), e25630.

Viscosi, V., Fortini, P., Slice, D., Loy, A., & Blasi, C. (2009). Geometric morphometric analyses of leaf variation in four oak species of the subgenus *Quercus* (Fagaceae). *Plant Biosystems, 143*(3), 575-587.

Vogler, A. P., & Monaghan, M. T. (2007). Recent advances in DNA taxonomy. *Journal of Zoological Systematics and Evolutionary Research, 45*(1), 1-10. Retrieved from https://doi.org/10.1111/j.1439-0469.2006.00384.x. doi:10.1111/j.1439-0469.2006.00384.x

Vranckx, G., Jacquemyn, H., Muys, B., & Honnay, O. (2012). Meta‐analysis of susceptibility of woody plants to loss of genetic diversity through habitat fragmentation. *Conservation Biology, 26*(2), 228-237. Retrieved from https://conbio.onlinelibrary.wiley.com/doi/abs/10.1111/j.1523-1739.2011.01778.x.

Wäldchen, J., & Mäder, P. (2018). Plant species identification using computer vision techniques: A systematic literature review. *Archives of Computational Methods in Engineering, 25*(2), 507-543.

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology, 23*(23), 5649-5662. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.12938. doi:10.1111/mec.12938

Wang, R., Compton, S. G., & Chen, X.-Y. (2011). Fragmentation can increase spatial genetic structure without decreasing pollen-mediated gene flow in a wind-pollinated tree. *Molecular Ecology, 20*(21), 4421-4432. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2011.05293.x. doi:10.1111/j.1365-294X.2011.05293.x

Watanabe, A. (2018). How many landmarks are enough to characterize shape and size variation?. *PloS One*, *13*(6), e0198341.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F‐statistics for the analysis of population structure. *Evolution, 38*(6), 1358-1370.

Wen, J., Lu, L. M., & Boggan, J. K. (2013). Diversity and evolution of Vitaceae in the Philippines. *Philippine Journal of Science*, *142*(Special Issue), 223-244.

Westoby, J. (1989). *Introduction to world forestry: people and their trees*: Basil Blackwell.

White, T. J., Bruns, T., Lee, S. J. W. T., & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications*, *18*(1), 315-322.

Wright, S. (1978). *Evolution and the genetics of populations: a treatise in four volumes: Vol. 4: variability within and among natural populations*. University of Chicago Press.

Xu, F., Guo, W., Xu, W., Wei, Y., & Wang, R. (2009). Leaf morphology correlates with water and light availability: what consequences for simple and compound leaves? *Progress in Natural Science, 19*(12), 1789-1798.

Yang, L., Kong, H., Huang, J.-P., & Kang, M. (2019). Different species or genetically divergent populations? Integrative species delimitation of the *Primulina hochiensis* complex from isolated karst habitats. *Molecular Phylogenetics and Evolution, 132*, 219-231. Retrieved from http://www.sciencedirect.com/science/article/pii/S1055790318302732. doi:https://doi.org/10.1016/j.ympev.2018.12.011

Young, A., Boyle, T., & Brown, T. (1996). The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution, 11*(10), 413-418.

Zakaria, W., Ahmad Puad, A., Zainudin, R., & Latiff, A. (2017). A revision of *Tetrastigma* (Miq.) Planch. (Vitaceae) in Sarawak, Borneo. *Malayan Nature Journal, 69*, 71-90.

Zakaria, W., Fatiha, W. N., Puad, A., Shafreena, A., Geri, C., Zainudin, R., & Latiff, A. (2016). *Tetrastigma diepenhorstii* (Miq.) Latiff (Vitaceae), a new host of *Rafflesia tuan-mudae* Becc.(Rafflesiaceae) in Borneo. *Journal of Botany, 2016*. Article ID 3952323.

Zelditch, M. L., Swiderski, D. L., & Sheets, H. D. (2004). *Geometric morphometrics for biologists: a primer*: Academic Press.

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics, 29*(22), 2869-2876. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3810850/pdf/btt499.pdf.

APPENDICES

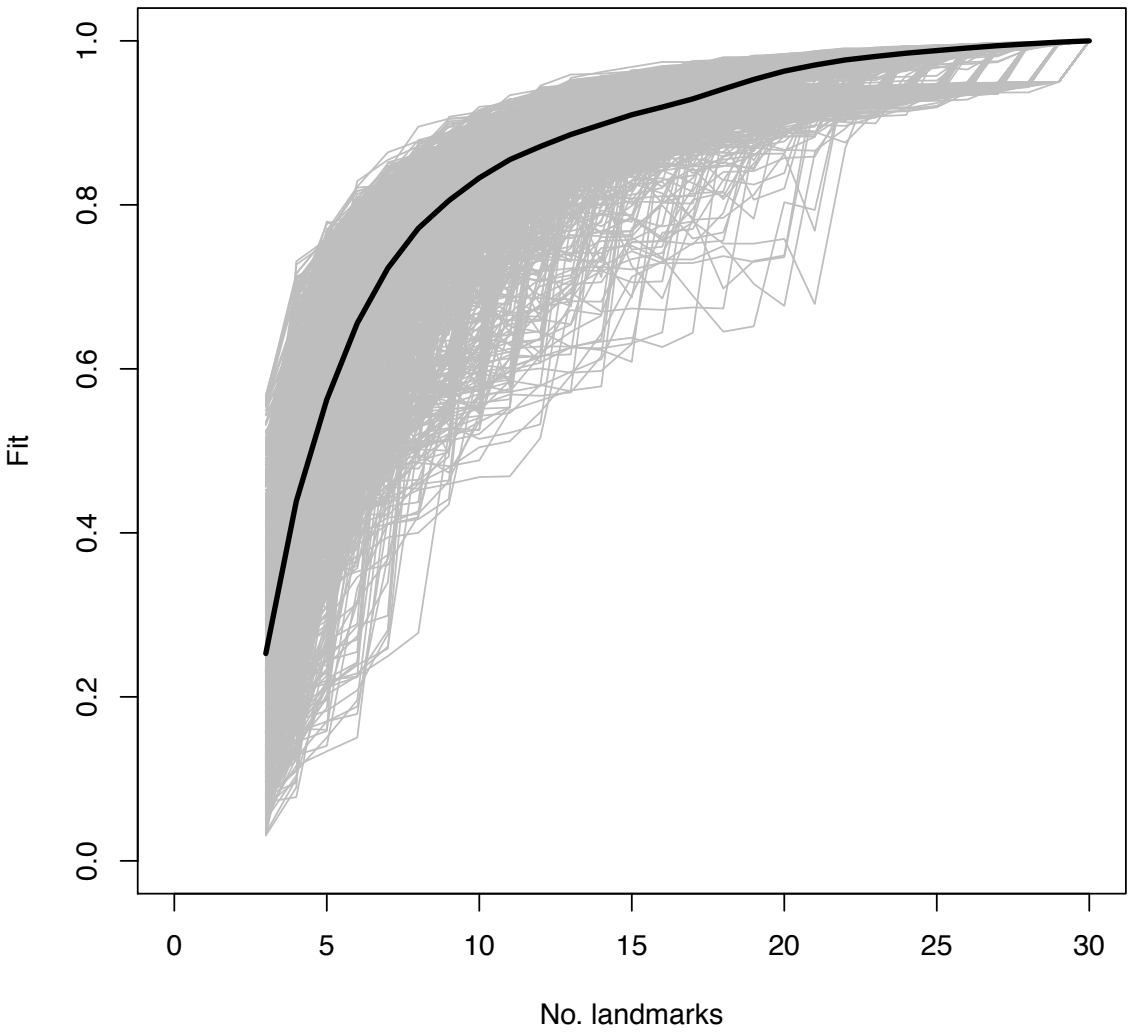
Appendix 1. Specimens of Tetrastigma loheri s.l. used in this study. The island region refers to the three major island regions in the Philippines (i.e. Luzon, Visayas, Mindanao). Voucher location: CAHUP- University of the Philippines Los Baños in Laguna, CANU- University of Canterbury in New Zealand, CEBU- University of San Carlos in Cebu, PNH - National Museum of the Philippines in Manila City, PUH- University of the Philippines in Quezon City. Samples are sorted according to collecting number.

Collection number	Species	Province	Island region	Voucher location
Barcelona 3476 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3478 Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3483 Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3487 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3491 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3495 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3498 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3501 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3503 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3505 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3507 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3509 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3518 with Nickrent & Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU/CAHUP
Barcelona 3554 with D.L. Nickrent, Pelser, & Bagacina	Tetrastigma loheri s.l.	Camarines Sur	Luzon	CANU/CAHUP
Barcelona 3565 with Co	Tetrastigma loheri s.l.	Bataan	Luzon	CANU/CAHUP
Barcelona 3569 with Co	Tetrastigma loheri s.l.	Bataan	Luzon	CANU/CAHUP
Barcelona 3573 with Co	Tetrastigma loheri s.l.	Bataan	Luzon	CANU/CAHUP
Barcelona 3581 with Co	Tetrastigma loheri s.l.	Camarines Sur	Luzon	CANU/CAHUP
Barcelona 3598 with Payba, Echanique, & Tabuc	Tetrastigma loheri s.l.	Cagayan	Luzon	CANU/CAHUP
Barcelona 3601 with Payba, Echanique, & Tabuc	Tetrastigma loheri s.l.	Cagayan	Luzon	CANU/CAHUP
Barcelona 3604 with Payba, Echanique, & Tabuc	Tetrastigma loheri s.l.	Cagayan	Luzon	CANU/CAHUP
Barcelona 3606 with Payba, Echanique, & Tabuc	Tetrastigma loheri s.l.	Cagayan	Luzon	CANU/CAHUP
Barcelona 3608 with Payba, Echanique, & Tabuc	Tetrastigma loheri s.l.	Cagayan	Luzon	CANU/CAHUP
Barcelona 3612 with Payba, Echanique, & Tabuc	Tetrastigma loheri s.l.	Cagayan	Luzon	CANU/CAHUP
Barcelona 3613 with Payba, Echanique, & Tabuc	Tetrastigma loheri s.l.	Cagayan	Luzon	CANU/CAHUP
Barcelona 3615 with Payba, Echanique, & Tabuc	Tetrastigma loheri s.l.	Cagayan	Luzon	CANU/CAHUP
Barcelona 3618 with Payba, Gabriel, & Aresta	Tetrastigma loheri s.l.	Cagayan	Luzon	CANU/CAHUP
Barcelona 3622 with Payba, Ocampo, & Vidad	Tetrastigma loheri s.l.	Cagayan	Luzon	CANU/CAHUP
Barcelona 3636 with Nickrent & Badilla	Tetrastigma loheri s.l.	Compostela	Mindanao	CANU/CAHUP

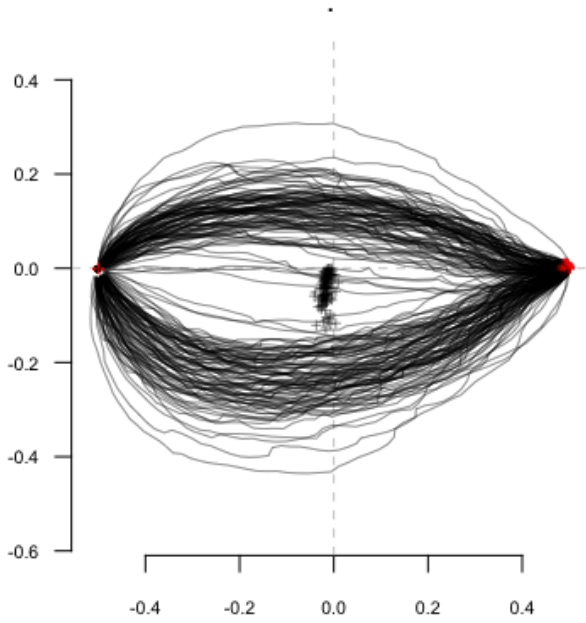
Collection number	Species	Province	Island region	Voucher location
Barcelona 3660 with Sarmiento	Tetrastigma loheri s.l.	Quezon	Luzon	CANU/CAHUP
Barcelona 3667 with Sarmiento	Tetrastigma loheri s.l.	Quezon	Luzon	CANU/CAHUP
Barcelona 3676 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU
Barcelona 3680 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU
Barcelona 3684 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU
Barcelona 3694 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU
Barcelona 3708 with Pelser	Tetrastigma loheri s.l.	Iloilo	Visayas	CANU
Barcelona 3718 with Pelser	Tetrastigma loheri s.l.	Iloilo	VIsayas	CANU
Barcelona 3726 et al.	Tetrastigma loheri s.l.	Samar	Visayas	CANU
Barcelona 3747 with Pelser & Gapas	Tetrastigma loheri s.l.	Camarines Sur	Luzon	CANU
Barcelona 3748 with Pelser & Gapas	Tetrastigma loheri s.l.	Camarines Sur	Luzon	CANU
Barcelona 3750 with Pelser	Tetrastigma loheri s.l.	Albay	Luzon	CANU
Barcelona 3751 with Pelser	Tetrastigma loheri s.l.	Albay	Luzon	CANU
Barcelona 3756 et al.	Tetrastigma loheri s.l.	Aurora	Luzon	CANU/PNH
Barcelona 3763 et al.	Tetrastigma loheri s.l.	Aurora	Luzon	CANU/PNH
Barcelona 3771 with Pelser	Tetrastigma loheri s.l.	Antique	Visayas	CANU
Barcelona 3818 with Pelser	Tetrastigma loheri s.l.	Quezon	Luzon	CANU
Barcelona 3826 with Pelser	Tetrastigma loheri s.l.	Quezon	Luzon	CANU
Barcelona 3827 with Pelser	Tetrastigma loheri s.l.	Camarines Norte	Luzon	CANU
Barcelona 3831 with Pelser	Tetrastigma loheri s.l.	Camarines Norte	Luzon	CANU
Barcelona 3851 with Pelser	Tetrastigma loheri s.l.	Camarines Norte	Luzon	CANU
Barcelona 3862 with Pelser	Tetrastigma loheri s.l.	Nuevea Ecija	Luzon	CANU
Barcelona 3868 with Pelser	Tetrastigma loheri s.l.	Nuevea Ecija	Luzon	CANU
Barcelona 3907 with Pelser	Tetrastigma loheri s.l.	Antique	VIsayas	CANU
Barcelona 4007 with Pelser	Tetrastigma loheri s.l.	Negros Occidental	Visayas	CANU
Barcelona 4038	Tetrastigma loheri s.l.	Aurora	Luzon	CANU
Barcelona 4050	Tetrastigma loheri s.l.	Camarines Sur	Luzon	CANU
Barcelona 4053	Tetrastigma loheri s.l.	Albay	Luzon	PNH
Barcelona 4056	Tetrastigma loheri s.l.	Nuevea Ecija	Luzon	CANU/PNH
Barcelona 4059	Tetrastigma loheri s.l.	Nuevea Ecija	Luzon	CANU/PNH
Barcelona 4062	Tetrastigma loheri s.l.	Nueva Ecija	Luzon	CANU/PNH
Barcelona 4066	Tetrastigma loheri s.l.	Nueva Ecija	Luzon	CANU/PNH
Barcelona 4067	Tetrastigma loheri s.l.	Nueva Ecija	Luzon	CANU/PNH
Barcelona 4075	Tetrastigma loheri s.l.	Nueva Ecija	Luzon	CANU/PNH
Barcelona 4077	Tetrastigma loheri s.l.	Nuevea Ecija	Luzon	CANU/PNH
Barcelona 4079	Tetrastigma loheri s.l.	Nueva Ecija	Luzon	CANU/PNH
Barcelona 4082	Tetrastigma loheri s.l.	Nuevea Ecija	Luzon	CANU/PNH

Collection number	Species	Province	Island region	Voucher location
Barcelona 4085	Tetrastigma loheri s.l.	Nueva Ecija	Luzon	CANU/PNH
Barcelona 4087	Tetrastigma loheri s.l.	Nueva Ecija	Luzon	CANU/PNH
Barcelona 4090 with Pelser	Tetrastigma loheri s.l.	Ilocos Norte	Luzon	CANU/PNH
Barcelona 4105 with Pelser	Tetrastigma loheri s.l.	Surigao	Mindanao	CANU/PNH
Barcelona 4106 with Pelser	Tetrastigma loheri s.l.	Surigao	Mindanao	CANU/PNH
Barcelona 4107 with Pelser	Tetrastigma loheri s.l.	Surigao	Mindanao	CANU/PNH
Barcelona 4144 with Pelser	Tetrastigma loheri s.l.	Iloilo	Visayas	CANU/PNH
Barcelona 4156 with Pelser	Tetrastigma loheri s.l.	Iloilo	Visayas	CANU/PNH
Barcelona 4172 with Pelser	Tetrastigma loheri s.l.	Negros Occidental	Visayas	CANU
Barcelona 4174 with Pelser	Tetrastigma loheri s.l.	Cebu	Visayas	CANU
Callado 390	Tetrastigma loheri s.l.	Aurora	Luzon	CANU/PNH
Callado 392	Tetrastigma loheri s.l.	Aurora	Luzon	CANU/PNH
Callado 394	Tetrastigma loheri s.l.	Aurora	Luzon	CANU/PNH
Obico 300	Tetrastigma cf. loheri	Cebu	Visayas	CANU, CEBU
Obico 304	Tetrastigma cf. loheri	Cebu	Visayas	CANU, CEBU
Obico 354	Tetrastigma cf. loheri	Cebu	Visayas	CANU, CEBU
Obico 368	Tetrastigma cf. loheri	Cebu	Visayas	CANU, CEBU
Obico 409	Tetrastigma cf. loheri	Cebu	Visayas	CANU
Obico 427	Tetrastigma cf. loheri	Cebu	Visayas	CANU, CEBU
Obico 453	Tetrastigma cf. loheri	Cebu	Visayas	CANU
Obico 476	Tetrastigma cf. loheri	Cebu	Visayas	CANU
Obico 516	Tetrastigma cf. loheri	Cebu	Visayas	CANU, CEBU
Obico 574	Tetrastigma cf. loheri	Cebu	Visayas	CANU
Obico 900	Tetrastigma cf. loheri	Bataan	Luzon	CANU
Obico 903	Tetrastigma cf. stenophyllum	La Union	Luzon	CANU, PUH
Obico 904	Tetrastigma cf. philippinense	Benguet	Luzon	CANU, PUH
Obico 972	Tetrastigma cf. loheri	Cebu	Visayas	CANU, CEBU
Obico 983	Tetrastigma cf. trifoliolatum	Leyte	Visayas	CANU, PUH
Obico 984	Tetrastigma cf. trifoliolatum	Leyte	Visayas	CANU, PUH
Obico 986	Tetrastigma cf. loheri	Leyte	Visayas	CANU, PUH
Obico 992	Tetrastigma cf. loheri	Leyte	Visayas	CANU

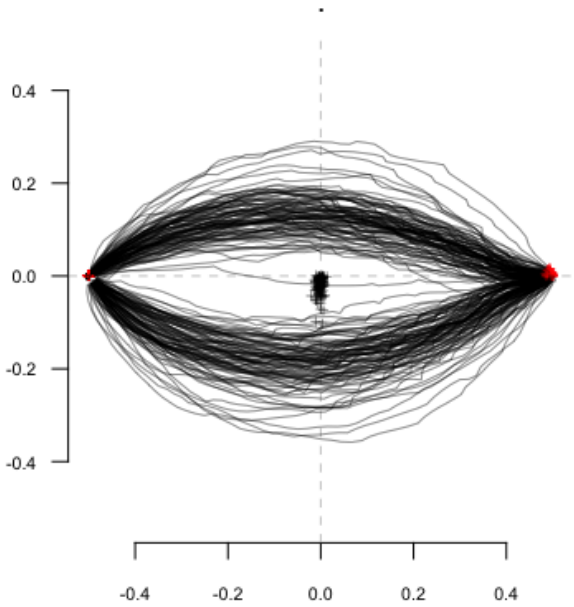
Appendix 2. A sampling curve from a lasec analysis of 30 landmarks used on 97 individuals of Tetrastigma loheri s.l.



Appendix 3. Aligned and superimposed lateral lamina outlines of 97 individuals of Tetrastigma loheri s.l.



Appendix 4. Aligned and superimposed terminal lamina outlines of 97 individuals of Tetrastigma loheri s.l.



Appendix 5. Momocs R script used in the analysis of the leaf outline of lateral and terminal lamina of 97 individuals of Tetrastigma loheri s.l.

Momocs script

```
library(dplyr)
library(purrr)
library(tidyverse)
library(Momocs)

#import photos and extract coordinates from the photos

lat<-import_jpg() #lateral leaves
ter<-import_jpg() #terminal leaves
?import_jpg

#create list of filenames and assign it an object called list
#filename should include the name of samples and the different levels it belongs to

lat_names<-list.files(path='/Users/Jasper/Pictures/momocs jpg/L_binary')
ter_names<-list.files(path='/Users/Jasper/Pictures/momocs jpg/T_binary')

#telling lf_structure that there are 6 groups in the filename
lfac<-lf_structure(lat_names, names=character(6), split='_', trim.extension=T)
tfac<-lf_structure(ter_names, names=character(6), split='_', trim.extension=T)

#putting names to columns

colnames(lfac)<-c("lineage", "species","major_region","island","island2","individual")
names(lfac)
lfac

colnames(tfac)<-c("lineage","species","major_region","island","island2","individual")
names(tfac)
tfac

#extract the list
write.table(lat_names, file="lat_list.txt", sep="\t")
write.table(ter_names, file="ter_list.txt",sep="\t")

#creates outlines and converts coordinates into a Coo object
##lateral leaves
#fac refers to factor aka the categories or groupings
?Out
l_out<-Out(lat,fac=lfac)
l_out
panel(l_out)
panel(l_out,fac="island2",names=F)


##terminal leaves
t_out<-Out(ter,fac=tfac)
t_out$ldk
panel(t_out)
panel(t_out,fac="major_region",names=T)


#defining landmarks
##lateral leaves

lat_lm<-def_ldk(l_out, 2)
lat_lm
save(lat_lm, file="final_lat.rda")
getwd()

##terminal leaves

ter_lm<-def_ldk(t_out, 2)
ter_lm
save(ter_lm, file="final_ter.rda")
getwd()

#you can load the photos with landmarks that you saved by:
load("final_lat.rda")
```

```
load("final_ter.rda")

#Alignment and Scaling
##lateral leaves

lat_lm %>% panel
x <- lat_lm %>% coo_bookstein %>% coo_slide(ldk=1) %>% arrange(individual)
Latareas <- x %>% coo_slice(ldk=1:2) %>% map(coo_area)

Lat_flip_flag <- Latareas[[1]]>Latareas[[2]]

Lat_flipx_subset <- function(x, cond){
  combine(filter(x,  cond) %>% coo_flipx() %>% coo_rev() %>% mutate(flip=TRUE),
          filter(x, !cond) %>% mutate(flip=FALSE))
}

xlat_ready <- combine(
  x %>% Lat_flipx_subset(Lat_flip_flag),
  x %>% Lat_flipx_subset(Lat_flip_flag)
)
xlat_ready$ldk %<>% map(~c(1, .x[2]))

xlat_ready %>% stack
save(xlat_ready, file="xlat_ready.rda")
xf_l <- xlat_ready %>% coo_bookstein() %>% coo_sample(360) %>% efourier(norm=FALSE)

#PCA
lat_pca<-xf_l %>% PCA()
plot_PCA(lat_pca,~island,chull=F,morphospace = T)%>%layer_points(pch = 16, cex=1)%>%layer_labelgroups(rect=F,
cex=0.8)%>%layer_ellipses(conf=0.9) #%>%layer_legend() # #%>%layer_ellipses(conf=0.9)%>%
plot_PCA(lat_pca,~island2,chull=F,morphospace = T)%>%layer_labelgroups(rect=F)%>%layer_points(pch = 16, cex=1)
#%>%layer_ellipses(conf=0.9)
?plot_PCA


####terminal leaves
ter_lm %>% panel
y <- ter_lm %>% coo_bookstein %>% coo_slide(ldk=1) %>% arrange(individual)
Terareas <- y %>% coo_slice(ldk=1:2) %>% map(coo_area)

Ter_flip_flag <- Terareas[[1]]>Terareas[[2]]

Ter_flipx_subset <- function(y, cond){
  combine(filter(y,  cond) %>% coo_flipx() %>% coo_rev() %>% mutate(flip=TRUE),
          filter(y, !cond) %>% mutate(flip=FALSE))
}

yter_ready <- combine(
  y %>% Ter_flipx_subset(Ter_flip_flag),
  y %>% Ter_flipx_subset(Ter_flip_flag)
)
yter_ready$ldk %<>% map(~c(1, .x[2]))

yter_ready %>% stack
yf_t <- yter_ready %>% coo_bookstein() %>% coo_sample(360) %>% efourier(norm=FALSE)


#PCA
ter_pca<-yf_t %>% PCA()
ter_pca%>% plot_PCA(~island,chull=F)%>%layer_labelgroups(rect=F)%>%layer_points(pch = 16,
cex=1)%>%layer_ellipses(conf=0.9)


#######combine two laminas
#arrange the names first
arrange(xf_l,individual)%>%export(,file="try_lat.txt") #checking them in print
arrange(yf_t,individual)%>%export(,file="try_ter.txt")

larr<-arrange(xf_l,individual)
tarr<-arrange(yf_t,individual)

?combine
comb_shapearr<-combine(tarr, larr) #this resolved the mislabeling of one specimen
comb_shapearr
export(comb_shapearr,file="rename_combshape.txt")
```

```
getwd()
comb_pca<-comb_shapearr %>% PCA()
plot_PCA(comb_pca,~island,chull=F,morphospace = T)%>%layer_labelgroups(rect=F)%>%layer_points(pch =
16,cex=1)%>%layer_ellipses(conf=0.9)

plot_PCA(comb_pca,~major_region,chull=F,morphospace = T)%>%layer_labelgroups(rect=F)%>%layer_points(pch = 21)
%>%layer_ellipses(conf=0.9)
```

Appendix 6. Collected specimens of Tetrastigma loheri *s.l. and other* Tetrastigma *species and their respective voucher locations. Samples are sorted according to collecting number.*

Collecting number	Species ID	Voucher data	Voucher location
Barcelona 3483 with Pelser	Tetrastigma loheri s.l.	Panay Island, Antique	CANU/CAHUP
Barcelona 3492 with Pelser	Ampelocissus sp.	Panay Island, Antique	CANU/CAHUP
Barcelona 3493 with Pelser	Tetrastigma papillosum	Panay Island, Antique	CANU/CAHUP
Barcelona 3539 with Nickrent & Pelser	Tetrastigma harmandii	Panay Island, Antique	CANU/CAHUP
Barcelona 3541 with Nickrent & Pelser	Tetrastigma cf. magnum	Luzon, Camarines Sur Province, Mt. Isarog	CANU/CAHUP
Barcelona 3562 with Co	Tetrastigma sp. A	Luzon, Bataan, Mt. Natib	CANU/CAHUP
Barcelona 3571 with Co	Tetrastigma ellipticum s.l.	Luzon, Bataan, Mt. Natib	CANU/CAHUP
Barcelona 3573 with Co	Tetrastigma loheri s.l.	Luzon, Bataan, Mt. Natib	CANU/CAHUP
Barcelona 3581 with Co	Tetrastigma loheri s.l.	Luzon, Camarines Sur, Mt. Labo	CANU/CAHUP
Barcelona 3592 with Payba, Echanique, & Tabuc	Tetrastigma sp. A	Luzon, Cagayan	CANU/CAHUP
Barcelona 3598 with Payba, Echanique, & Tabuc	Tetrastigma loheri s.l.	Luzon, Cagayan	CANU/CAHUP
Barcelona 3601 with Payba, Echanique, & Tabuc	Tetrastigma loheri s.l.	Luzon, Cagayan	CANU/CAHUP
Barcelona 3615 with Payba, Echanique, & Tabuc	Tetrastigma loheri s.l.	Luzon, Cagayan	CANU/CAHUP
Barcelona 3618 with Payba, Gabriel, & Aresta	Tetrastigma loheri s.l.	Luzon, Cagayan	CANU/CAHUP
Barcelona 3635 with Nickrent & Badilla	Tetrastigma cf. magnum	Mindanao, Compostela Valley	CANU/CAHUP
Barcelona 3636 with Nickrent & Badilla	Tetrastigma loheri s.l.	Mindanao, Compostela Valley	CANU/CAHUP
Barcelona 3667 with Sarmiento	Tetrastigma loheri s.l.	Luzon, Quezon, Mt. Irid	CANU/CAHUP
Barcelona 3680 with Pelser	Tetrastigma loheri s.l.	Panay Island, Antique	CANU
Barcelona 3708 with Pelser	Tetrastigma loheri s.l.	Panay Island, Iloilo	CANU
Barcelona 3718 with Pelser	Tetrastigma loheri s.l.	Panay Island, Iloilo	CANU
Barcelona 3726 et al.	Tetrastigma loheri s.l.	Samar, Basey	CANU
Barcelona 3746 with Pelser and Gapas	Cayratia sp.	Luzon, Camarines Sur, Buhi	CANU
Barcelona 3748 with Pelser & Gapas	Tetrastigma loheri s.l.	Luzon, Camarines Sur	CANU
Barcelona 3749 with Pelser	Cayratia sp.	Luzon, Camarines Sur, Mt. Malinao	CANU
Barcelona 3751 with Pelser	Tetrastigma loheri s.l.	Luzon, Malinao	CANU
Barcelona 3754 with Pelser	Tetrastigma ellipticum s.l.	Luzon, Aurora	Photos only
Barcelona 3756 et al.	Tetrastigma loheri s.l.	Luzon, Aurora	CANU/PNH
Barcelona 3763 et al.	Tetrastigma loheri s.l.	Luzon, Aurora	CANU/PNH
Barcelona 3765 with Pelser	Cayratia sp.	Panay Island, Iloilo	CANU
Barcelona 3766 with Pelser	Tetrastigma harmandii	Panay, Antique	CANU
Barcelona 3771 with Pelser	Tetrastigma loheri s.l.	Panay Island, Antique	CANU
Barcelona 3778 with Pelser	Tetrastigma papillosum	Mindanao, Bukidnon	Photos only
Barcelona 3818 with Pelser	Tetrastigma loheri s.l.	Luzon, Quezon, Mt. Banahaw	CANU
Barcelona 3825 with Pelser	Cayratia sp.	Luzon, Quezon, Mt. Banahaw	CANU
Barcelona 3826 with Pelser	Tetrastigma loheri s.l.	Luzon, Quezon, Mt. Banahaw	CANU
Barcelona 3827 with Pelser	Tetrastigma loheri s.l.	Luzon, Camarines Norte	CANU
Barcelona 3851 with Pelser	Tetrastigma loheri s.l.	Luzon, Camarines Norte	No information
Barcelona 3868 with Pelser	Tetrastigma loheri s.l.	Luzon, Nueva Ecija, Mt. Mingan	CANU
Barcelona 3904 with Pelser	Tetrastigma cf. scariosum	Luzon, Nueva Ecija, Mt. Mingan	CANU
Barcelona 3907 with Pelser	Tetrastigma loheri s.l.	Panay, Antique	CANU
Barcelona 4007 with Pelser	Tetrastigma loheri s.l.	Negros, Mt. Kanlaon Natural Park	CANU
Barcelona 4038	Tetrastigma loheri s.l.	Luzon, Aurora	CANU
Barcelona 4053	Tetrastigma loheri s.l.	Luzon, Albay	PNH
Barcelona 4054	Causonis trifolia	Luzon, Palawan	no voucher
Barcelona 4059	Tetrastigma loheri s.l.	Luzon, Nueva Ecija	CANU/PNH
Barcelona 4062	Tetrastigma loheri s.l.	Luzon, Nueva Ecija	CANU/PNH
Barcelona 4066	Tetrastigma loheri s.l.	Luzon, Nueva Ecija	CANU/PNH
Barcelona 4067	Tetrastigma loheri s.l.	Luzon, Nueva Ecija	CANU/PNH
Barcelona 4075	Tetrastigma loheri s.l.	Luzon, Nueva Ecija	CANU/PNH
Barcelona 4082	Tetrastigma loheri s.l.	Luzon, Nueva Ecija	CANU/PNH
Barcelona 4085	Tetrastigma loheri s.l.	Luzon, Nueva Ecija	CANU/PNH

Collecting number	Species ID	Voucher data	Voucher location
Barcelona 4087	Tetrastigma loheri s.l.	Luzon, Nueva Ecija	CANU/PNH
Barcelona 4090 with Pelser	Tetrastigma loheri s.l.	Luzon, Ilocos Norte	CANU/PNH
Barcelona 4092 with Pelser	Tetrastigma sp. A	Luzon, Ilocos Norte	CANU/PNH
Barcelona 4101 with Pelser	Tetrastigma aff. glabratum	Mindanao, Surigao del Sur	CANU/PNH
Barcelona 4105 with Pelser	Tetrastigma loheri s.l.	Mindanao, Surigao del Sur	CANU/PNH
Barcelona 4106 with Pelser	Tetrastigma loheri s.l.	Mindanao, Surigao del Sur	CANU/PNH
Barcelona 4107 with Pelser	Tetrastigma loheri s.l.	Mindanao, Surigao del Sur	CANU/PNH
Barcelona 4108 with Pelser	Tetrastigma cf. scariosum	Mindanao, Surigao del Sur	CANU/PNH
Barcelona 4144 with Pelser	Tetrastigma loheri s.l.	Panay, Iloilo	CANU/PNH
Barcelona 4156 with Pelser	Tetrastigma loheri s.l.	Panay, Iloilo	CANU/PNH
Barcelona 4193	Tetrastigma loheri s.l.	Cebu, Argao	No voucher
Barcelona 4195	Tetrastigma loheri s.l.	Cebu, Argao	No voucher
Callado 1159	Tetrastigma aff. glabratum	Mindanao, Bukidnon	PNH
Callado 1230	Tetrastigma aff. glabratum	Mindanao, Bukidnon	PNH
Callado 390	Tetrastigma loheri s.l.	Luzon, Aurora	CANU/PNH
Nickrent 5531 with Van Ee & Barcelona	Tetrastigma aff. glabratum	Mindanao, Bukidnon	No information
Obico 300	Tetrastigma loheri s.l.	Cebu, Tabunan	CANU, CEBU
Obico 355	Tetrastigma loheri s.l.	Cebu, Balamban	No voucher
Obico 361	Tetrastigma loheri s.l.	Cebu, Balamban	No voucher
Obico 371	Tetrastigma loheri s.l.	Cebu, Balamban	No voucher
Obico 386	Tetrastigma loheri s.l.	Cebu, Balamban	No voucher
Obico 453	Tetrastigma loheri s.l.	Cebu, Argao	CANU
Obico 469	Tetrastigma loheri s.l.	Cebu, Dalaguete	No voucher
Obico 473	Tetrastigma loheri s.l.	Cebu, Dalaguete	CANU
Obico 575	Tetrastigma loheri s.l.	Cebu, Boljoon	No voucher
Obico 580	Tetrastigma loheri s.l.	Cebu, Boljoon	No voucher
Obico 608	Tetrastigma loheri s.l.	Cebu, Boljoon,	No voucher
Obico 900	Tetrastigma loheri s.l.	Luzon, Bataan	CANU
Obico 903	Tetrastigma stenophyllum	Luzon, La Union	CANU/PUH
Obico 904	Tetrastigma philippinense	Luzon, Benguet	CANU/PUH
Obico 972	Tetrastigma loheri s.l.	Cebu, Tabunan	CANU/CEBU
Obico 983	Tetrastigma cf. trifoliolatum	Leyte, Dagami	CANU/PUH
Obico 984	Tetrastigma cf. trifoliolatum	Leyte, Dagami	CANU/PUH
Obico 986	Tetrastigma loheri s.l.	Leyte, Dagami	CANU/PUH
Obico 992	Tetrastigma loheri s.l.	Leyte, Dagami	CANU

Appendix 7. Primers and their respective sequences for amplification of the nuclear (ITS and ETS) and plastid regions (atpB-rbcL , psbA-trnH , rps16 , trnL, and trnL-F).

DNA regions	Primers	Reference	Sequence (5'–3')
atpb-rbcl	atpB2	Manen et al. (1994)	GAAGTAGTAGGATTGATTCTC
	rbcL5	Manen et al. (1994)	TACAGTTGTCCATGTACCAG
rps16	rpsF	Oxelman et al. (1997)	GTGGTAGAAAGCAACGTGCGACTT
	rpsR2	Oxelman et al. (1997)	TCGGGATCGAACATCAATTGCAAC
ITS	ITS Rosetto F	Rossetto et al. (2002)	CCTGCGGAAGGATCATTG
	ITS 4	White et al. (1990)	TCCTCCGCTTATTGATATGC
	ITS A	Blattner (1999)	GGAAGGAGAAGTCGTAACAAGG
	ITS C	Blattner (1999)	GCAATTCACACCAAGTATCGC
	ITS E	Blattner (1999)	CGGCAACGGATATCTCGGCTC
	ITS B	Blattner (1999)	CTTTTCCTCCGCTTATTGATATG
psba	psbA	Hamilton (1999)	GTTATGCATGAACGTAATGCTC
	trnH	Hamilton (1999)	CGCGCATGGTGGATTCACAAATC
trnL-F	tab E	Taberlet et al. (1991)	GGTTCAAGTCCCTCTATCCC
	tab F	Taberlet et al. (1991)	ATTTGAACTGGTGACACGAG
trnL	tab C	Taberlet et al. (1991)	CGAAATCGGTAGACGCTACG
	tab D	Taberlet et al. (1991)	GGGGATAGAGGGACTTGAAC
ETS	3F	This study	GTTGGCAGGCTCCTTGCTTA
	4F	This study	TTGGCAGGCTCCTTGCTTAT
	5F	This study	TGGCAGGCTCCTTGCTTATG
	614F	This study	CAGCTCTAGAATTACTACGGTTATCC
	615R	This study	TCAGCTCTAGAATTACTACGGTTATCC
	84F	This study	GTTGGCAGGCTCCTTGCTTA
	85F	This study	TTGGCAGGCTCCTTGCTTAT
	86F	This study	TGGCAGGCTCCTTGCTTATG
	87F	This study	GGCAGGCTCCTTGCTTATGT
	699R	This study	TGGCACGTATCAGCTCTAGA
	ETS forward	Okuyama et al. (2004)	GGTGCCTAAAATGCGTGGGTGGACAGG
	18S-IGS	Baldwin and Markos (1998)	GAGACAAGCATATGACTACTGGCAGGATCAACCAG
	18S-E	Baldwin and Markos (1998)	GCAGGATCAACCAGGTAGCA
	18S-ETS	Baldwin and Markos (1998)	ACTTACACATGCATGGCTTAATCT

Appendix 8. Accessions of Vitis vinifera and species from Saxifragaceae used to design primers for the External Transcribed Spacer region for Tetrastigma.

Accession	GenBank code
Bensoniella oregona	AB291995.1
Conimitella williamsii	AB291996.1
Elmera racemosa	AB248766.1
Heuchera rubescens	AB292005.1
Heuchera merriamii	KJ828787.1
Lithophragma bolanderi	AB292008.1
Mitella stylosa var. makinoi,	AB163444.1
Mitella kiusiana	AB163443.1
Tiarella cordifolia	AB292014.1
Tolmiea menziesiii	AB248774.1
Vitis vinifera	AM463304.2

Appendix 9. Specimens with corresponding species names and their Genbank accessions used in this study. New sequences generated in this study will be submitted to GenBank after a manuscript from this chapter is accepted for publication. NA refers to missing data. The total number of accessions per specimen and per DNA region is provided in the last column and row respectively.

Specimen	Species name	ETS	ITS	atpb-rbcL	psbA-trnH	rps16	trnL and trnL-F	Total
Ampelociss2821UniPan	Ampelocissus sp. Barcelona 3492	This study	NA	KT597333	KT597460	KT597280	KT597216	5
Causonis_tri3233UniTay	Causonis trifolia	NA	KT597084	KT597334	KT597461	KT597281	KT597217	5
Cayratia_clematidea	Cayratia clematidea	NA	NA	KC166297	NA	KC166388	KC166625	3
Cayratia2529UniMal	Cayratia sp. Barcelona 3749	NA	NA	KT597335	KT597462	KT597282	KT597218	4
Cayratia2542UniBuh	Cayratia sp. Barcelona 3746	NA	NA	KT597336	NA	KT597283	KT597219	3
Cayratia2949UniIlo	Cayratia sp. Barcelona 3765	NA	NA	KT597337	NA	NA	NA	1
Cayratia3059UniBan	Cayratia sp. Barcelona 3825	NA	NA	KT597338	NA	NA	NA	1
T_aff_glabratumJO3290Min	Tetrastigma aff. glabratum	This study	KT597209	KT597453	NA	NA	NA	3
T_cf_loheri_JO3654_Argao	Tetrastigma loheri s.l.	This study	This study	NA	NA	NA	NA	2
T_cf_loheri_Leyte_JO4026	Tetrastigma loheri s.l.	This study	This study	This study	NA	This study	This study	5
T_cf_loheri2419Min	Tetrastigma loheri s.l.	This study	This study	KT597339	NA	NA	NA	3
T_cf_loheri2423Cag	Tetrastigma loheri s.l.	This study	KT597085	KT597340	NA	NA	NA	3
T_cf_loheri2424Cag	Tetrastigma loheri s.l.	This study	KT597086	KT597341	KT597463	KT597284	KT597220	6
T_cf_loheri2428Cag	Tetrastigma loheri s.l.	This study	KT597087	KT597342	KT597464	KT597285	KT597221	6
T_cf_loheri2431Cag	Tetrastigma loheri s.l.	NA	KT597088	KT597343	KT597465	KT597286	KT597222	5
T_cf_loheri2437CSr	Tetrastigma loheri s.l.	This study	KT597089	KT597344	KT597466	NA	KT597223	5
T_cf_loheri2447Bat	Tetrastigma loheri s.l.	NA	KT597090	NA	KT597467	KT597287	KT597224	4
T_cf_loheri2539Mal	Tetrastigma loheri s.l.	This study	KT597091	KT597345	NA	NA	NA	3
T_cf_loheri2546Sam	Tetrastigma loheri s.l.	NA	KT597092	KT597346	KT597468	KT597288	KT597225	5
T_cf_loheri2550CSr	Tetrastigma loheri s.l.	NA	KT597093	NA	NA	NA	NA	1
T_cf_loheri2700Aur	Tetrastigma loheri s.l.	This study	KT597094	KT597347	KT597469	KT597289	KT597226	6
T_cf_loheri2701Aur	Tetrastigma loheri s.l.	This study	KT597095	KT597348	NA	NA	NA	3

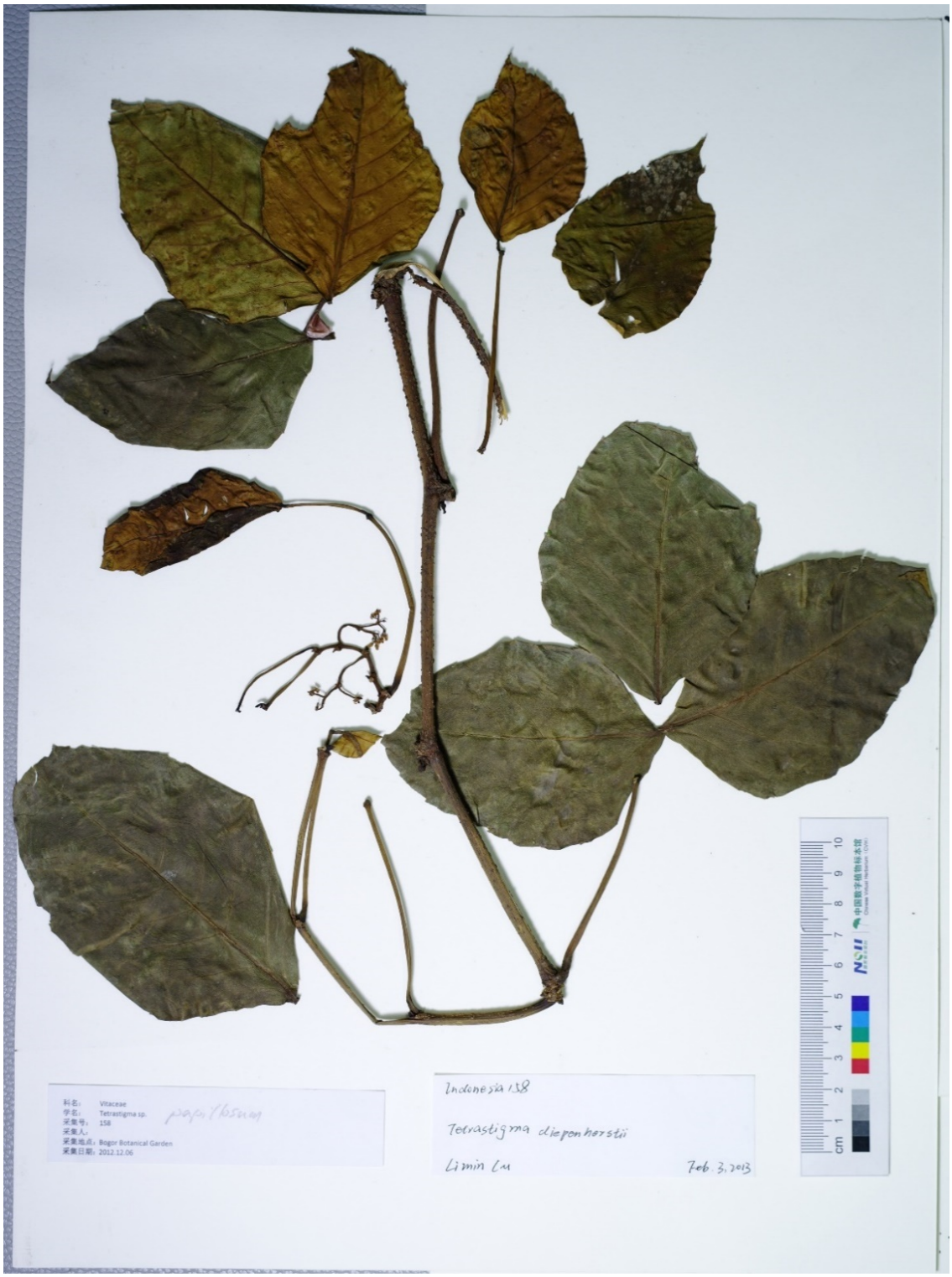
Specimen	Species name	ETS	ITS	atpb-rbcL	psbA-trnH	rps16	trnL and trnL-F	Total
T_cf_loheri2702Aur	Tetrastigma loheri s.l.	NA	KT597096	KT597349	NA	NA	NA	2
T_cf_loheri2787Que	Tetrastigma loheri s.l.	This study	KT597097	NA	KT597470	NA	KT597227	4
T_cf_loheri2789Pan	Tetrastigma loheri s.l.	This study	KT597098	KT597350	KT597471	KT597290	KT597228	6
T_cf_loheri2791Pan	Tetrastigma loheri s.l.	This study	KT597099	KT597351	KT597472	KT597291	KT597229	6
T_cf_loheri2795Pan	Tetrastigma loheri s.l.	NA	KT597100	KT597352	KT597473	KT597292	KT597230	5
T_cf_loheri2828Pan	Tetrastigma loheri s.l.	This study	KT597101	KT597353	NA	KT597293	KT597231	5
T_cf_loheri2955Pan	Tetrastigma loheri s.l.	This study	KT597102	KT597354	NA	NA	NA	3
T_cf_loheri3060Que	Tetrastigma loheri s.l.	NA	KT597103	KT597355	NA	NA	NA	2
T_cf_loheri3199Neg	Tetrastigma loheri s.l.	This study	KT597104	NA	NA	NA	NA	2
T_cf_loheri3240Aur	Tetrastigma loheri s.l.	This study	KT597105	NA	NA	NA	NA	2
T_cf_loheri3248NEc	Tetrastigma loheri s.l.	NA	KT597106	NA	NA	NA	NA	1
T_cf_loheri3251NEc	Tetrastigma loheri s.l.	NA	KT597107	KT597356	NA	NA	NA	2
T_cf_loheri3255NEc	Tetrastigma loheri s.l.	NA	KT597108	NA	NA	NA	NA	1
T_cf_loheri3256NEc	Tetrastigma loheri s.l.	This study	KT597109	KT597357	NA	NA	NA	3
T_cf_loheri3264NEc	Tetrastigma loheri s.l.	NA	KT597110	NA	NA	NA	NA	1
T_cf_loheri3271NEc	Tetrastigma loheri s.l.	This study	KT597111	KT597358	NA	NA	NA	3
T_cf_loheri3274NEc	Tetrastigma loheri s.l.	This study	KT597112	KT597359	NA	NA	NA	3
T_cf_loheri3276ManCar	Tetrastigma loheri s.l.	NA	KT597113	KT597360	NA	NA	NA	2
T_cf_loheri3276NEc	Tetrastigma loheri s.l.	This study	NA	NA	NA	NA	NA	1
T_cf_loheri3279INr	Tetrastigma loheri s.l.	This study	KT597114	KT597361	NA	NA	NA	3
T_cf_loheri3294Min	Tetrastigma loheri s.l.	This study	KT597115	KT597362	NA	NA	NA	3
T_cf_loheri3314Pan	Tetrastigma loheri s.l.	This study	KT597118	KT597365	NA	NA	NA	3
T_cf_loheri3324Pan	Tetrastigma loheri s.l.	This study	KT597119	KT597366	NA	NA	NA	3
T_cf_loheriJO4033Ley	Tetrastigma loheri s.l.	This study	This study	This study	NA	This study	This study	5
T_cf_magnum2420Mir	Tetrastigma cf. magnum	This study	KT597136	KT597383	KT597486	KT597303	KT597244	6
T_cf_trifoliolatum_Leyte_JO4018	Tetrastigma cf. trifoliolatum	This study	This study	NA	This study	This study	NA	4

Specimen	Species name	ETS	ITS	atpb-rbcL	psbA-trnH	rps16	trnL and trnL-F	Total
T_cf_trifoliolatum_Leyte_JO4022	Tetrastigma cf. trifoliolatum	This study	This study	NA	NA	This study	NA	3
T_cf_tuberculatumWen10280	Tetrastigma cf. tuberculatum	NA	NA	HM585559	HM585699	HM585835	HM585975	4
T_crenatum	Tetrastigma crenatum	NA	AY037909	NA	NA	NA	AF300313	2
T_diepenhorstii_ChenLu158	Tetrastigma diepenhorstii	NA	NA	KY766323	KY766775	KY766661	KY766833	4
T_diepenhorstii8261	Tetrastigma diepenhorstii	NA	NA	HM585567	NA	HM585843	HM585983	3
T_ellipticum2450UniNat	Tetrastigma ellipticum s.l.	This study	KT597144	KT597391	NA	NA	NA	3
T_ellipticumJO2524	Tetrastigma ellipticum s.l.	This study	KT597154	KT597401	KT597495	KT597310	KT597253	6
T_harmandii2950UniEga	Tetrastigma harmandii	This study	KT597165	KT597412	NA	NA	NA	3
T_harmandiiJO2395	Tetrastigma harmandii	This study	KT597128	KT597375	KT597480	NA	KT597238	5
T_lawsonii7503	Tetrastigma lawsonii	NA	NA	HM585598	HM585736	HM585873	HM586014	4
T_lawsonii7505	Tetrastigma lawsonii	NA	NA	HM585599	HM585737	HM585874	HM586015	4
T_loheri_Bataan_JO3932	Tetrastigma loheri s.l.	This study	This study	NA	NA	This study	This study	4
T_loheri_JO3502Tab	Tetrastigma loheri s.l.	This study	This study	NA	NA	NA	NA	2
T_loheri_JO3557Bal	Tetrastigma loheri s.l.	This study	This study	NA	NA	NA	NA	2
T_loheri_JO3563Bal	Tetrastigma loheri s.l.	This study	This study	NA	NA	NA	NA	2
T_loheri_JO3572Bal	Tetrastigma loheri s.l.	This study	This study	NA	NA	NA	NA	2
T_loheri_JO3670Dal	Tetrastigma loheri s.l.	This study	This study	NA	NA	NA	NA	2
T_loheri_JO3674Dal	Tetrastigma loheri s.l.	This study	This study	NA	NA	NA	NA	2
T_loheri_JO3727Alc	Tetrastigma loheri s.l.	This study	This study	NA	NA	NA	NA	2
T_loheri_JO3729Alc	Tetrastigma loheri s.l.	This study	This study	NA	NA	NA	NA	2
T_loheri_JO3756Blj	Tetrastigma loheri s.l.	This study	This study	NA	NA	NA	NA	2
T_loheri_JO4007Tab	Tetrastigma loheri s.l.	This study	This study	This study	This study	This study	This study	6
T_loheri3053Que	Tetrastigma loheri s.l.	This study	KT597121	KT597368	KT597477	KT597297	KT597235	6
T_loheri3061CNr	Tetrastigma loheri s.l.	This study	KT597122	KT597369	NA	NA	NA	3
T_loheri3080CNr	Tetrastigma loheri s.l.	NA	KT597123	KT597370	NA	NA	NA	2
T_loheri3139Pan	Tetrastigma loheri s.l.	This study	KT597125	KT597372	NA	NA	NA	3

Specimen	Species name	ETS	ITS	atpb-rbcL	psbA-trnH	rps16	trnL and trnL-F	Total
T_loheri3232Alb	Tetrastigma loheri s.l.	This study	KT597126	KT597373	NA	NA	NA	3
T_loheri3481Arg	Tetrastigma loheri s.l.	NA	This study	NA	NA	NA	NA	1
T_loheri3483Arg	Tetrastigma loheri s.l.	NA	This study	NA	NA	NA	NA	1
T_loheriJO3134NEc	Tetrastigma loheri s.l.	This study	KT597124	KT597371	KT597478	KT597298	KT597236	6
T_magnum2394	Tetrastigma cf. magnum	This study	KT597127	KT597374	KT597479	KT597299	KT597237	6
T_papillosumJO2805Pan	Tetrastigma papillosum	This study	KT597162	KT597409	KT597500	KT597315	KT597258	6
T_papillosumJO2962Min	Tetrastigma papillosum	This study	KT597170	KT597417	NA	NA	NA	3
T_philippinenseJO3938Ben	Tetrastigma loheri s.l.	This study	This study	This study	This study	This study	This study	6
T_spA2433	Tetrastigma sp. A	This study	KT597139	KT597386	NA	NA	NA	3
T_spA2455	Tetrastigma sp. A	NA	KT597147	KT597394	KT597491	KT597306	KT597249	5
T_spA3281UniPag	Tetrastigma sp. A	This study	KT597204	KT597448	NA	NA	NA	3
T_stenophyllum_LaUnion_JO3936	Tetrastigma stenophyllum	This study	This study	This study	This study	This study	This study	6
T_trifoliolatum10758	Tetrastigma trifoliolatum	NA	NA	HM585643	HM585780	HM585916	HM586058	4
T_trifoliolatum8350	Tetrastigma trifoliolatum	NA	NA	HM585644	HM585781	HM585917	HM586059	4
T_tuberculatumWen6668	Tetrastigma tuberculatum	NA	NA	HM585649	HM585784	HM585922	HM586064	4
T_tuberculatumWen8335	Tetrastigma tuberculatum	NA	NA	HM585651	HM585785	HM585924	HM586066	4
Tetrastigma_aff_glabratum_2475	Tetrastigma aff. glabratum	This study	KT597151	KT597398	KT597492	KT597307	KT597250	6
Tetrastigma_aff_glabratum_2965_Min	Tetrastigma aff. glabratum	NA	KT597171	KT597418	NA	NA	NA	2
Tetrastigma_aff_glabratum_2968_Min	Tetrastigma aff. glabratum	NA	KT597174	KT597420	NA	NA	NA	2
Tetrastigma_cf_loheri_3295	Tetrastigma loheri s.l.	This study	KT597116	KT597363	KT597474	KT597294	KT597232	6
Tetrastigma_cf_loheri_3296	Tetrastigma loheri s.l.	This study	KT597117	KT597364	KT597475	KT597295	KT597233	6
Tetrastigma_cf_scariosum_3137	Tetrastigma cf. scariosum	This study	KT597187	KT597432	KT597512	KT597324	KT597270	6
Tetrastigma_cf_scariosum_3297	Tetrastigma cf. scariosum	NA	KT597210	KT597454	KT597518	KT597330	KT597277	5
Tetrastigma_glabratum_Wen_10666	Tetrastigma glabratum	NA	NA	HM585579	NA	HM585855	HM585995	3
Tetrastigma_nitens	Tetrastigma nitens	NA	AF365984	KY766342	KY766796	KY766682	EF179093	5
Tetrastigma_petraeum	Tetrastigma petraeum	NA	AY037910	NA	NA	NA	EF179094	2

Specimen	Species name	ETS	ITS	atpb-rbcL	psbA-trnH	rps16	trnL and trnL-F	Total
Tetrastigma_sp_Wen_10768	Tetrastigma sp.	NA	NA	HM585637	HM585774	HM585910	HM586052	4
Tetrastigma_sp_Wen_8256	Tetrastigma sp.	NA	NA	HM585638	HM585775	HM585911	HM586053	4
Tetrastigma_sp_Wen_8455	Tetrastigma sp.	NA	NA	HM585653	HM585787	HM585926	HM586068	4
Tetrastigma_strumarum_Wen_10757	Tetrastigma strumarum	NA	NA	HM585641	HM585778	HM585914	HM586056	4
	Total	62	83	78	44	50	53	370

Appendix 10. Tetrastigma diepenhorstii (Chen & Lu 158; PE) from Habib et al., 2017.



Appendix 11.General morphology of Tetrastigma loheri.

A. Leaf and inflorescence.

B. Woody habit and fruits.

C. Young red shoots.

Appendix 12. Geographic coordinates and herbarium repository of specimens of Tetrastigma loheri *collected from 4 forested areas on Cebu island. CCPL- Central Cebu Protected Landscape. CANU - University of Canterbury Herbarium, New Zealand. CEBU- University of San Carlos Herbarium, Cebu, Philippines. The table is sorted by forest area.*

Sample number	Collecting number	Forest area	Site	Geographic coordinates	Herbarium repository
3726	Obico 574	Alcoy	Upper Becerril	9° 38' 34.728'' N, 123° 25' 25.428'' E	CANU
3727	Obico 575	Alcoy	Upper Becerril	9° 38' 34.188'' N, 123° 25' 23.484'' E	No voucher
3729	Obico 580	Alcoy	Upper Becerril	9° 37' 53.6'' N, 123° 26' 5'' E	No voucher
3750	Obico 602	Alcoy	Nangka	9° 40' 57.0'' N, 123° 25' 14.1'' E	No voucher
3752	Obico 604	Alcoy	San Antonio	9° 40' 51.5'' N, 123° 25' 1.9'' E	No voucher
3753	Obico 605	Alcoy	San Antonio	9° 40' 16.3'' N, 123° 24' 44.4'' E	No voucher
3754	Obico 606	Alcoy	San Antonio	9° 40' 10.1'' N, 123° 24' 45'' E	No voucher
3755	Obico 607	Alcoy	San Antonio	9° 39' 34.0'' N, 123° 24' 45.4'' E	No voucher
3756	Obico 608	Alcoy	San Antonio	9° 40' 20.6'' N, 123° 25' 1.7'' E	No voucher
3757	Obico 609	Alcoy	San Antonio	9° 40' 18.5'' N, 123° 25' 2.9'' E	No voucher
3769	Obico 621	Alcoy	San Antonio	9° 40' 10.1'' N, 123° 24' 59.4'' E	No voucher
3774	Obico 643	Alcoy	Nug-as	9° 43' 25.608'' N, 123° 26' 56.688'' E	No voucher
3778	Obico 673	Alcoy	Nug-as	9° 43' 4.872'' N, 123° 26' 51.432'' E	No voucher
3779	Obico 678	Alcoy	Nug-as	9° 43' 4.008'' N, 123° 26' 49.920'' E	No voucher
3785	Obico 685	Alcoy	Nug-as	9° 42' 58.068'' N, 123° 26' 48.084'' E	No voucher
3787	Obico 695	Alcoy	Nug-as	9° 42' 54.540'' N, 123° 26' 45.996'' E	No voucher
3790	Obico 698	Alcoy	Nug-as	9° 42' 53.568'' N, 123° 26' 45.996'' E	No voucher
3798	Obico 708	Alcoy	Nug-as	9° 42' 51.624'' N, 123° 26' 47.220'' E	No voucher
3807	Obico 719	Alcoy	Nug-as	9° 42' 48.528'' N, 123° 26' 48.912'' E	No voucher
3817	Obico 733	Alcoy	Nug-as	9° 43' 17.112'' N, 123° 26' 59.388'' E	No voucher
3836	Obico 774	Alcoy	Nug-as	9° 42' 57.780'' N, 123° 27' 3.024'' E	No voucher
3843	Obico 790	Alcoy	Nug-as	9° 42' 58.176'' N, 123° 27' 12.888'' E	No voucher
3847	Obico 799	Alcoy	Nug-as	9° 43' 0.264'' N, 123° 27' 18.540'' E	No voucher
3849	Obico 805	Alcoy	Nug-as	9° 42' 33.7'' N, 123° 25' 12'' E	No voucher
3850	Obico 808	Alcoy	Nug-as	9° 43' 30.1'' N, 123° 26' 34.8'' E	No voucher
3851	Obico 809	Alcoy	Nug-as	9° 43' 42.2'' N, 123° 26' 43.5'' E	No voucher
3856	Obico 818	Alcoy	Poblacion	9° 43' 4.1'' N, 123° 28' 00'' E	No voucher
3481	Barcelona 4193	Argao	Cansuje	9° 55' 39.1'' N, 123° 31' 21.3'' E	No voucher
3482	Barcelona 4194	Argao	Cansuje	9° 55' 28.8'' N, 123° 30' 43.6'' E	CANU
3483	Barcelona 4195	Argao	Cansuje	9° 56' 19.5'' N, 123° 30' 43.9'' E	No voucher
3610	Obico 409	Argao	Cansuje	9° 55' 12.540'' N, 123° 32' 6.936'' E	CANU, CEBU
3612	Obico 411	Argao	Cansuje	9° 55' 1.776'' N, 123° 31' 6.924'' E	No voucher
3613	Obico 412	Argao	Cansuje	9° 54' 56.952'' N, 123° 30' 56.700'' E	No voucher
3614	Obico 413	Argao	Cansuje	9° 54' 54.612'' N, 123° 30' 54.036'' E	No voucher
3615	Obico 414	Argao	Cansuje	9° 54' 51.228'' N, 123° 30' 48.528'' E	No voucher
3617	Obico 416	Argao	Cansuje	9° 54' 53.892'' N, 123° 30' 45.072'' E	No voucher

Sample number	Collecting number	Forest area	Site	Geographic coordinates	Herbarium repository
3618	Obico 417	Argao	Cansuje	9° 54' 56.016'' N, 123° 30' 43.920'' E	No voucher
3619	Obico 418	Argao	Cansuje	9° 54' 56.376'' N, 123° 30' 42.300'' E	No voucher
3625	Obico 424	Argao	Cansuje	9° 55' 0.048'' N, 123° 30' 38.772'' E	No voucher
3626	Obico 425	Argao	Cansuje	9° 55' 6.924'' N, 123° 30' 36.468'' E	No voucher
3627	Obico 426	Argao	Cansuje	9° 56' 11.436'' N, 123° 31' 11.964'' E	No voucher
3628	Obico 427	Argao	Cansuje	9° 56' 13.812'' N, 123° 31' 7.788'' E	CANU
3629	Obico 428	Argao	Cansuje	9° 56' 12.732'' N, 123° 31' 5.952'' E	No voucher
3630	Obico 429	Argao	Cansuje	9° 56' 12.552'' N, 123° 31' 4.656'' E	No voucher
3631	Obico 430	Argao	Cansuje	9° 56' 12.228'' N, 123° 31' 3.936'' E	No voucher
3632	Obico 431	Argao	Cansuje	9° 56' 10.824'' N, 123° 31' 1.812'' E	No voucher
3633	Obico 432	Argao	Cansuje	9° 56' 10.032'' N, 123° 30' 58.752'' E	No voucher
3635	Obico 434	Argao	Cansuje	9° 56' 11.904'' N, 123° 30' 57.816'' E	CANU
3636	Obico 435	Argao	Cansuje	9° 56' 15.612'' N, 123° 30' 58.680'' E	No voucher
3637	Obico 436	Argao	Cansuje	9° 56' 17.880'' N, 123° 31' 0.336'' E	CANU
3639	Obico 438	Argao	Cansuje	9° 54' 10.224'' N, 123° 31' 47.208'' E	No voucher
3640	Obico 439	Argao	Cansuje	9° 54' 10.044'' N, 123° 31' 46.056'' E	No voucher
3644	Obico 443	Argao	Cansuje	9° 54' 8.640'' N, 123° 31' 41.340'' E	No voucher
3645	Obico 444	Argao	Tabayag	9° 54' 8.496'' N, 123° 31' 40.476'' E	CANU
3646	Obico 445	Argao	Tabayag	9° 54' 8.424'' N, 123° 31' 41.088'' E	CANU
3647	Obico 446	Argao	Tabayag	9° 53' 16.764'' N, 123° 31' 30.324'' E	No voucher
3648	Obico 447	Argao	Tabayag	9° 59' 20.724'' N, 123° 30' 47.916'' E	No voucher
3649	Obico 448	Argao	Tabayag	9° 53' 24.504'' N, 123° 30' 45.900'' E	No voucher
3650	Obico 449	Argao	Tabayag	9° 53' 21.264'' N, 123° 30' 43.920'' E	No voucher
3651	Obico 450	Argao	Canbantug	9° 53' 16.224'' N, 123° 30' 40.824'' E	No voucher
3652	Obico 451	Argao	Canbantug	9° 53' 13.848'' N, 123° 30' 41.112'' E	No voucher
3653	Obico 452	Argao	Canbantug	9° 53' 11.328'' N, 123° 30' 40.392'' E	No voucher
3654	Obico 453	Argao	Canbantug	9° 53' 8.304'' N, 123° 30' 39.240'' E	CANU
3655	Obico 454	Argao	Canbantug	9° 53' 5.424'' N, 123° 30' 35.856'' E	No voucher
3656	Obico 455	Argao	Canbantug	9° 53' 3.264'' N, 123° 30' 35.928'' E	No voucher
3658	Obico 457	Argao	Canbantug	9° 53' 0.456'' N, 123° 30' 36.144'' E	No voucher
3659	Obico 458	Argao	Canbantug	9° 52' 58.476'' N, 123° 30' 36.108'' E	No voucher
3661	Obico 460	Argao	Canbantug	9° 52' 53.184'' N, 123° 30' 33.156'' E	No voucher
3664	Obico 463	Argao	Canbantug	9° 52' 32.844'' N, 123° 30' 18.900'' E	No voucher
3665	Obico 464	Argao	Canbantug	9° 52' 34.536'' N, 123° 30' 18.936'' E	No voucher
3667	Obico 466	Argao	Canbantug	9° 55' 48.000'' N, 123° 30' 13.248'' E	No voucher
3668	Obico 467	Argao	Canbantug	9° 55' 44.580'' N, 123° 30' 15.120'' E	No voucher
3669	Obico 468	Argao	Canbantug	9° 55' 27.156'' N, 123° 30' 54.108'' E	No voucher
3502	Obico 300	CCPL	Tabunan	10° 26' 25.404'' N, 123° 49' 9.516'' E	CANU, CEBU
3503	Obico 301	CCPL	Tabunan	10° 26' 21.480'' N, 123° 49' 22.656'' E	No voucher
3504	Obico 302	CCPL	Tabunan	10° 26' 21.336'' N, 123° 49' 22.836'' E	No voucher

Sample number	Collecting number	Forest area	Site	Geographic coordinates	Herbarium repository
3505	Obico 303	CCPL	Tabunan	10° 26' 20.976'' N, 123° 49' 22.944'' E	No voucher
3506	Obico 304	CCPL	Tabunan	10° 26' 22.056'' N, 123° 49' 29.496'' E	CANU, CEBU
3507	Obico 305	CCPL	Tabunan	10° 26' 23.172'' N, 123° 49' 31.728'' E	No voucher
3513	Obico 311	CCPL	Tabunan	10° 26' 22.416'' N, 123° 49' 34.572'' E	CANU
3514	Obico 312	CCPL	Tabunan	10° 26' 22.416'' N, 123° 49' 34.608'' E	CANU, CEBU
3515	Obico 313	CCPL	Tabunan	10° 26' 32.604'' N, 123° 49' 21.972'' E	CANU, CEBU
3516	Obico 314	CCPL	Tabunan	10° 26' 32.604'' N, 123° 49' 23.160'' E	No voucher
3519	Obico 317	CCPL	Tabunan	10° 26' 30.228'' N, 123° 49' 32.412'' E	No voucher
3522	Obico 320	CCPL	Tabunan	10° 26' 29.832'' N, 123° 49' 32.988'' E	No voucher
3524	Obico 322	CCPL	Tabunan	10° 26' 29.688'' N, 123° 49' 33.060'' E	No voucher
3530	Obico 328	CCPL	Tabunan	10° 26' 28.536'' N, 123° 49' 35.004'' E	No voucher
3531	Obico 329	CCPL	Tabunan	10° 26' 28.248'' N, 123° 49' 35.112'' E	No voucher
3536	Obico 334	CCPL	Tabunan	10° 26' 27.528'' N, 123° 49' 36.624'' E	No voucher
3537	Obico 335	CCPL	Tabunan	10° 26' 26.808'' N, 123° 49' 37.488'' E	No voucher
3538	Obico 336	CCPL	Tabunan	10° 26' 26.520'' N, 123° 49' 37.560'' E	No voucher
3539	Obico 337	CCPL	Tabunan	10° 26' 26.196'' N, 123° 49' 37.668'' E	CANU, CEBU
3541	Obico 339	CCPL	Tabunan	10° 26' 25.008'' N, 123° 49' 37.776'' E	No voucher
3544	Obico 342	CCPL	Tabunan	10° 26' 20.616'' N, 123° 49' 37.200'' E	No voucher
3545	Obico 343	CCPL	Tabunan	10° 26' 33.144'' N, 123° 49' 41.556'' E	CANU
3547	Obico 345	CCPL	Tabunan	10° 26' 32.640'' N, 123° 49' 41.700'' E	No voucher
3551	Obico 349	CCPL	Tabunan	10° 26' 32.100'' N, 123° 49' 39.576'' E	No voucher
3553	Obico 351	CCPL	Tabunan	10° 26' 30.876'' N, 123° 49' 39.720'' E	No voucher
3557	Obico 355	CCPL	Putol	10° 27' 21.024'' N, 123° 49' 21.288'' E	No voucher
3560	Obico 358	CCPL	Putol	10° 27' 21.456'' N, 123° 49' 22.008'' E	No voucher
3561	Obico 359	CCPL	Putol	10° 27' 21.276'' N, 123° 49' 22.296'' E	No voucher
3563	Obico 361	CCPL	Putol	10° 27' 21.348'' N, 123° 49' 22.620'' E	No voucher
3564	Obico 362	CCPL	Putol	10° 27' 20.592'' N, 123° 49' 23.484'' E	No voucher
3567	Obico 365	CCPL	Putol	10° 27' 19.872'' N, 123° 49' 22.476'' E	No voucher
3569	Obico 368	CCPL	Putol	10° 27' 19.440'' N, 123° 49' 21.756'' E	CANU, CEBU
3571	Obico 370	CCPL	Mit-ol	10° 27' 27.540'' N, 123° 49' 6.384'' E	No voucher
3572	Obico 371	CCPL	Mit-ol	10° 27' 28.584'' N, 123° 49' 6.492'' E	No voucher
3587	Obico 386	CCPL	Mit-ol	10° 27' 43.200'' N, 123° 49' 17.256'' E	No voucher
3588	Obico 387	CCPL	Mit-ol	10° 27' 28.836'' N, 123° 49' 19.416'' E	No voucher
3589	Obico 388	CCPL	Mauyog	10° 27' 40.572'' N, 123° 49' 35.472'' E	No voucher
3590	Obico 389	CCPL	Mauyog	10° 27' 39.312'' N, 123° 49' 36.660'' E	No voucher
3592	Obico 391	CCPL	Mauyog	10° 27' 38.484'' N, 123° 49' 37.092'' E	CANU
3593	Obico 392	CCPL	Mauyog	10° 27' 38.556'' N, 123° 49' 37.488'' E	No voucher
3594	Obico 393	CCPL	Mauyog	10° 27' 36.468'' N, 123° 49' 40.908'' E	No voucher
3595	Obico 394	CCPL	Mauyog	10° 27' 32.004'' N, 123° 49' 41.772'' E	No voucher
3596	Obico 395	CCPL	Mauyog	10° 27' 34.488'' N, 123° 49' 42.060'' E	No voucher

Sample number	Collecting number	Forest area	Site	Geographic coordinates	Herbarium repository
3597	Obico 396	CCPL	Mauyog	10° 27' 37.260'' N, 123° 49' 42.708'' E	CANU
3603	Obico 402	CCPL	Mauyog	10° 27' 38.700'' N, 123° 49' 43.212'' E	No voucher
3606	Obico 405	CCPL	Mauyog	10° 27' 40.212'' N, 123° 49' 41.628'' E	No voucher
3608	Obico 407	CCPL	Mauyog	10° 27' 40.752'' N, 123° 49' 40.944'' E	No voucher
3484	Barcelona 4196	Dalaguete	Mag-alambac	9° 48' 48.5'' N, 123° 28' 11'' ° E	No voucher
3486	Barcelona 4198	Dalaguete	Mag-alambac	9° 48' 53.1'' N, 123° 28' 5.4'' E	No voucher
3487	Barcelona 4199	Dalaguete	Mag-alambac	9° 48' 54.5'' N, 123° 28' 6'' E	No voucher
3499	Barcelona 4211	Dalaguete	Mag-alambac	9° 48' 31.2'' N, 123° 28' 30.5'' E	No voucher
3670	Obico 469	Dalaguete	Mag-alambac	9° 53' 36.924'' N, 123° 28' 12.828'' E	No voucher
3671	Obico 470	Dalaguete	Mahangin	9° 48' 51.012'' N, 123° 28' 12.720'' E	No voucher
3672	Obico 471	Dalaguete	Mahangin	9° 48' 57.996'' N, 123° 28' 11.244'' E	No voucher
3674	Obico 473	Dalaguete	Mahangin	9° 48' 57.744'' N, 123° 28' 9.624'' E	CANU
3676	Obico 475	Dalaguete	Mahangin	9° 49' 2.640'' N, 123° 28' 12.000'' E	No voucher
3677	Obico 476	Dalaguete	Mahangin	9° 49' 1.848'' N, 123° 28' 12.720'' E	CANU
3678	Obico 477	Dalaguete	Cangcabalong	9° 48' 55.224'' N, 123° 28' 9.408'' E	No voucher
3681	Obico 480	Dalaguete	Cangcabalong	9° 48' 57.204'' N, 123° 28' 8.364'' E	No voucher
3682	Obico 481	Dalaguete	Cangcabalong	9° 48' 58.572'' N, 123° 28' 7.896'' E	No voucher
3685	Obico 484	Dalaguete	Cangcabalong	9° 48' 59.004'' N, 123° 28' 7.320'' E	No voucher
3687	Obico 492	Dalaguete	Mantalongon	9° 48' 29.736'' N, 123° 28' 31.836'' E	No voucher
3691	Obico 498	Dalaguete	Mantalongon	9° 48' 29.844'' N, 123° 28' 28.992'' E	No voucher
3693	Obico 503	Dalaguete	Mantalongon	9° 48' 21.456'' N, 123° 28' 27.084'' E	No voucher
3707	Obico 520	Dalaguete	Baybayun	9° 48' 52.668'' N, 123° 28' 23.016'' E	No voucher
3710	Obico 533	Dalaguete	Baybayun	9° 48' 52.704'' N, 123° 28' 20.568'' E	No voucher
3712	Obico 535	Dalaguete	Baybayun	9° 48' 51.336'' N, 123° 28' 20.352'' E	No voucher
3713	Obico 536	Dalaguete	Baybayun	9° 48' 50.652'' N, 123° 28' 19.920'' E	No voucher
3714	Obico 537	Dalaguete	Baybayun	9° 48' 50.364'' N, 123° 28' 19.740'' E	CANU
3715	Obico 539	Dalaguete	Baybayun	9° 48' 50.148'' N, 123° 28' 17.148'' E	CANU
3718	Obico 546	Dalaguete	Baybayun	9° 48' 49.500'' N, 123° 28' 14.988'' E	CANU
3721	Obico 560	Dalaguete	Baybayun	9° 48' 48.924'' N, 123° 28' 14.808'' E	CANU

Appendix 13. Scoring of 13 microsatellite loci for Tetrastigma loheri *in Cebu based on the fragment length size of their respective alleles. NA indicates missing data. Column labels: Sam-sample number of individuals, Site-collection localities, LOH-microsatellite loci of* T. loheri*, M-number of loci with missing data for each individual. Percentage of missing data for each locus and the M13 fluorescent primer used for each locus are given at the bottom.*

Sam	Site	LOH 931	LOH 931	LOH 1248	LOH 1248	LOH 936	LOH 936	LOH 684	LOH 684	LOH 412	LOH 412	LOH 663	LOH 663	LOH 749	LOH 749	LOH 868	LOH 868	LOH 1497	LOH 1497	LOH 688	LOH 688	LOH 505	LOH 505	LOH 865	LOH 865	LOH 1512	LOH 1512	M
3502	Tabunan	215	225	197	197	219	225	341	341	280	280	288	288	261	263	157	161	202	202	338	347	NA	NA	223	223	167	177	1
3503	Tabunan	229	243	197	201	217	223	341	341	283	283	306	306	231	269	149	195	NA	NA	323	323	NA	NA	239	241	167	167	2
3504	Tabunan	215	229	197	201	217	219	341	341	286	289	290	290	265	265	161	183	204	204	320	323	443	443	229	229	NA	NA	1
3505	Tabunan	225	247	197	197	219	219	341	341	265	265	306	306	NA	NA	161	197	NA	NA	338	344	NA	NA	239	239	165	167	3
3506	Tabunan	226	241	197	197	219	219	341	341	284	289	NA	NA	237	259	149	153	NA	NA	NA	NA	443	443	231	231	NA	NA	4
3507	Tabunan	225	243	197	197	219	219	341	341	271	289	290	290	243	259	NA	NA	204	204	320	323	443	443	237	237	NA	NA	2
3513	Tabunan	231	231	197	201	219	219	341	341	265	280	290	290	241	245	153	197	NA	NA	332	347	407	407	225	241	167	167	1
3514	Tabunan	231	231	197	201	219	219	341	341	265	280	290	290	241	245	195	197	200	204	332	347	407	407	225	241	167	167	0
3515	Tabunan	215	241	197	197	219	219	341	341	283	286	290	290	271	271	155	197	NA	NA	323	347	437	443	221	237	167	171	1
3516	Tabunan	215	243	197	197	209	219	341	344	283	286	290	306	253	253	155	197	202	204	323	323	407	413	245	245	167	171	0
3519	Tabunan	215	215	197	197	197	219	341	341	283	283	316	316	261	267	171	171	208	210	323	341	407	413	NA	NA	169	171	1
3522	Tabunan	207	211	187	197	219	219	341	344	250	274	302	302	243	249	173	193	204	208	323	326	422	443	225	225	169	177	0
3524	Tabunan	215	215	NA	NA	219	219	NA	NA	250	274	NA	NA	243	249	173	193	204	208	323	326	422	443	225	225	NA	NA	4
3530	Tabunan	205	215	187	197	209	217	341	344	271	289	NA	NA	NA	NA	173	173	200	204	320	320	449	449	241	241	171	173	2
3531	Tabunan	205	215	187	197	209	217	341	344	271	289	NA	NA	247	247	173	173	200	204	320	320	449	449	241	241	171	173	1
3536	Tabunan	229	231	197	201	219	219	341	341	277	283	318	320	249	253	181	183	204	210	314	347	443	443	237	237	173	173	0
3537	Tabunan	NA	NA	187	197	219	219	341	344	271	283	290	290	241	241	169	169	208	208	326	338	NA	NA	241	241	NA	NA	3
3538	Tabunan	215	243	197	201	219	219	341	341	262	277	290	290	243	243	169	169	204	208	323	323	443	443	225	231	165	177	0
3539	Tabunan	NA	NA	187	197	219	219	341	341	271	274	NA	NA	249	259	NA	NA	200	204	320	320	428	449	NA	NA	167	177	4
3541	Tabunan	229	247	187	197	219	219	341	344	265	271	306	306	NA	NA	151	197	NA	NA	338	338	443	449	NA	NA	167	167	3
3544	Tabunan	231	243	187	197	219	219	344	344	283	283	308	308	237	265	171	171	204	204	332	347	452	452	235	235	167	183	0
3545	Tabunan	229	231	187	197	219	219	341	341	271	274	290	290	249	253	NA	NA	204	210	332	338	449	449	243	243	165	167	1

Sam	Site	LOH 931	LOH 931	LOH 1248	LOH 1248	LOH 936	LOH 936	LOH 684	LOH 684	LOH 412	LOH 412	LOH 663	LOH 663	LOH 749	LOH 749	LOH 868	LOH 868	LOH 1497	LOH 1497	LOH 688	LOH 688	LOH 505	LOH 505	LOH 865	LOH 865	LOH 1512	LOH 1512	M
3547	Tabunan	225	247	197	201	217	219	341	341	286	289	290	318	235	245	197	197	204	210	338	341	NA	NA	241	241	NA	NA	2
3551	Tabunan	241	247	187	197	219	219	341	341	283	286	NA	NA	NA	NA	171	197	NA	NA	326	329	413	413	239	239	167	167	3
3553	Tabunan	247	247	197	197	219	219	341	341	265	283	316	318	245	255	161	161	204	208	320	326	443	443	NA	NA	167	167	1
3557	Putol	NA	NA	197	201	219	219	341	341	286	295	290	290	NA	NA	147	201	202	202	320	323	443	443	217	225	169	179	2
3560	Putol	NA	NA	197	203	197	219	341	341	280	295	290	290	NA	NA	147	197	200	202	323	323	443	443	225	225	165	169	2
3561	Putol	NA	NA	187	197	219	219	341	341	280	316	290	290	247	257	147	147	200	204	323	353	407	413	NA	NA	169	177	2
3563	Putol	243	243	197	201	219	219	341	341	253	295	NA	NA	243	247	157	157	202	204	323	326	NA	NA	241	241	167	167	2
3564	Putol	215	225	197	201	219	219	341	341	253	283	288	290	235	241	153	153	NA	NA	323	338	407	413	NA	NA	NA	NA	3
3567	Putol	215	215	187	197	197	217	341	341	265	271	288	290	243	247	183	195	202	204	338	341	443	443	241	241	167	171	0
3569	Putol	215	225	189	201	219	219	341	341	283	289	288	288	247	271	155	155	NA	NA	341	344	407	407	231	241	167	167	1
3571	Mit-ol	215	215	201	203	209	219	344	347	253	283	290	316	239	263	153	175	NA	NA	323	338	NA	NA	241	241	169	169	2
3572	Mit-ol	231	233	197	197	219	219	341	344	283	283	290	290	255	263	189	189	204	204	320	338	440	440	NA	NA	NA	NA	2
3587	Mit-ol	NA	NA	197	201	219	219	NA	NA	280	295	288	288	235	235	179	179	202	202	338	341	434	434	241	241	173	175	2
3588	Mit-ol	NA	NA	201	203	219	219	341	341	253	253	288	288	241	241	165	203	194	204	338	341	NA	NA	241	241	NA	NA	3
3589	Mauyog	NA	NA	197	201	219	219	341	341	253	253	NA	NA	229	245	171	171	202	204	320	344	443	443	241	241	NA	NA	3
3590	Mauyog	NA	NA	197	201	219	219	341	341	271	283	318	318	241	257	NA	NA	204	204	323	326	443	443	231	241	173	173	2
3592	Mauyog	215	241	201	209	219	219	341	341	NA	NA	316	318	251	251	161	165	204	210	323	338	NA	NA	241	241	NA	NA	3
3593	Mauyog	215	243	197	201	219	219	341	341	265	289	NA	NA	235	235	153	153	202	204	311	326	NA	NA	241	241	167	181	2
3594	Mauyog	215	243	187	197	217	219	341	341	265	271	NA	NA	NA	NA	NA	NA	200	204	323	341	443	443	NA	NA	167	167	4
3595	Mauyog	245	245	187	197	219	219	341	341	286	295	308	308	257	257	165	165	200	204	317	338	NA	NA	241	241	NA	NA	2
3596	Mauyog	NA	NA	197	201	209	217	341	341	238	265	288	288	249	261	NA	NA	200	200	341	344	NA	NA	241	241	167	189	3
3597	Mauyog	229	233	191	201	NA	NA	341	341	271	271	288	288	249	249	167	193	200	204	323	344	NA	NA	241	241	173	181	2
3603	Mauyog	243	243	197	201	217	219	341	344	265	271	290	290	NA	NA	165	203	194	204	323	347	NA	NA	241	241	NA	NA	3
3606	Mauyog	225	243	187	197	209	219	341	344	271	283	290	290	235	235	167	167	202	204	338	341	443	443	225	241	167	171	0
3608	Mauyog	225	243	187	197	219	219	NA	NA	265	271	290	290	NA	NA	167	213	204	204	341	344	NA	NA	227	243	171	189	3
3481	Cansuje	NA	NA	187	197	219	219	341	341	244	244	304	312	231	231	147	157	200	204	308	320	443	446	249	249	NA	NA	2

Sam	Site	LOH 931	LOH 931	LOH 1248	LOH 1248	LOH 936	LOH 936	LOH 684	LOH 684	LOH 412	LOH 412	LOH 663	LOH 663	LOH 749	LOH 749	LOH 868	LOH 868	LOH 1497	LOH 1497	LOH 688	LOH 688	LOH 505	LOH 505	LOH 865	LOH 865	LOH 1512	LOH 1512	M
3482	Cansuje	215	253	187	197	219	219	341	341	253	253	290	290	229	229	147	155	190	200	320	320	440	440	235	243	171	171	0
3483	Cansuje	227	227	189	199	219	219	341	341	247	247	308	308	NA	NA	203	203	200	208	320	323	443	443	239	249	169	171	1
3610	Cansuje	215	225	197	205	197	219	NA	NA	229	238	308	308	229	229	161	173	218	218	308	320	437	443	251	251	149	151	1
3612	Cansuje	227	263	187	197	219	219	341	341	247	247	310	312	NA	NA	NA	NA	200	204	320	329	437	437	245	245	145	163	2
3613	Cansuje	227	241	187	197	219	219	344	344	235	247	310	310	253	253	NA	NA	208	218	320	320	443	443	243	243	169	171	1
3614	Cansuje	227	227	187	197	219	219	341	344	NA	NA	290	290	NA	NA	163	163	200	202	314	323	440	440	NA	NA	171	171	3
3615	Cansuje	243	259	197	197	219	219	341	341	247	247	306	308	229	241	151	213	200	202	323	323	404	440	243	243	171	173	0
3617	Cansuje	241	243	197	201	219	219	341	344	247	247	288	290	221	233	177	177	200	202	320	323	440	440	243	243	171	171	0
3618	Cansuje	227	241	197	197	219	219	341	344	NA	NA	290	310	221	233	NA	NA	200	202	320	326	440	440	229	243	149	171	2
3619	Cansuje	227	243	197	197	219	219	344	344	244	244	288	290	221	241	181	181	208	218	326	326	404	434	241	243	169	171	0
3625	Cansuje	227	227	197	197	219	219	341	344	244	247	290	290	233	233	163	163	200	200	314	323	440	440	241	241	169	175	0
3626	Cansuje	227	243	197	203	219	219	341	344	NA	NA	306	310	NA	NA	163	171	200	218	323	326	440	440	241	241	NA	NA	3
3627	Cansuje	227	243	197	199	219	219	341	341	247	247	290	310	NA	NA	155	169	190	200	326	341	434	434	249	249	NA	NA	2
3628	Cansuje	215	227	201	201	219	219	341	341	NA	NA	290	304	217	229	183	183	200	204	320	323	NA	NA	239	239	149	173	2
3629	Cansuje	231	243	189	199	197	197	341	344	247	247	290	308	NA	NA	151	159	190	200	311	323	440	440	239	239	167	171	1
3630	Cansuje	227	227	197	197	219	219	341	341	NA	NA	306	308	NA	NA	183	201	200	204	NA	NA	NA	NA	233	249	167	171	4
3631	Cansuje	241	241	197	203	219	219	341	341	244	247	290	310	NA	NA	167	167	200	200	317	320	443	443	NA	NA	NA	NA	3
3632	Cansuje	241	241	197	203	219	219	341	341	247	247	290	290	231	231	167	167	200	200	317	320	NA	NA	NA	NA	171	173	2
3633	Cansuje	243	243	197	197	217	219	341	341	NA	NA	290	308	229	229	159	159	200	204	314	314	NA	NA	NA	NA	NA	NA	4
3635	Cansuje	243	243	187	197	219	219	341	344	NA	NA	308	308	NA	NA	175	177	NA	NA	NA	NA	437	437	243	249	NA	NA	5
3636	Cansuje	NA	NA	187	197	217	219	341	344	247	247	290	310	217	217	NA	NA	200	208	320	320	437	437	NA	NA	171	171	3
3637	Cansuje	221	227	197	199	217	219	341	341	247	247	304	332	NA	NA	151	159	204	208	NA	NA	437	443	241	241	167	171	2
3667	Cansuje	243	243	197	199	219	219	341	341	247	247	290	306	NA	NA	171	209	190	200	320	338	NA	NA	229	229	169	169	2
3668	Cansuje	NA	NA	187	197	209	219	341	344	244	244	288	304	NA	NA	157	165	NA	NA	314	326	443	443	247	247	169	169	3
3669	Cansuje	205	215	187	197	219	219	341	344	247	265	304	304	NA	NA	191	191	200	204	323	326	440	440	249	249	145	145	1
3639	Tabayag	227	241	193	203	219	219	NA	NA	244	247	290	310	209	221	154	163	204	208	323	341	437	443	243	243	147	163	1

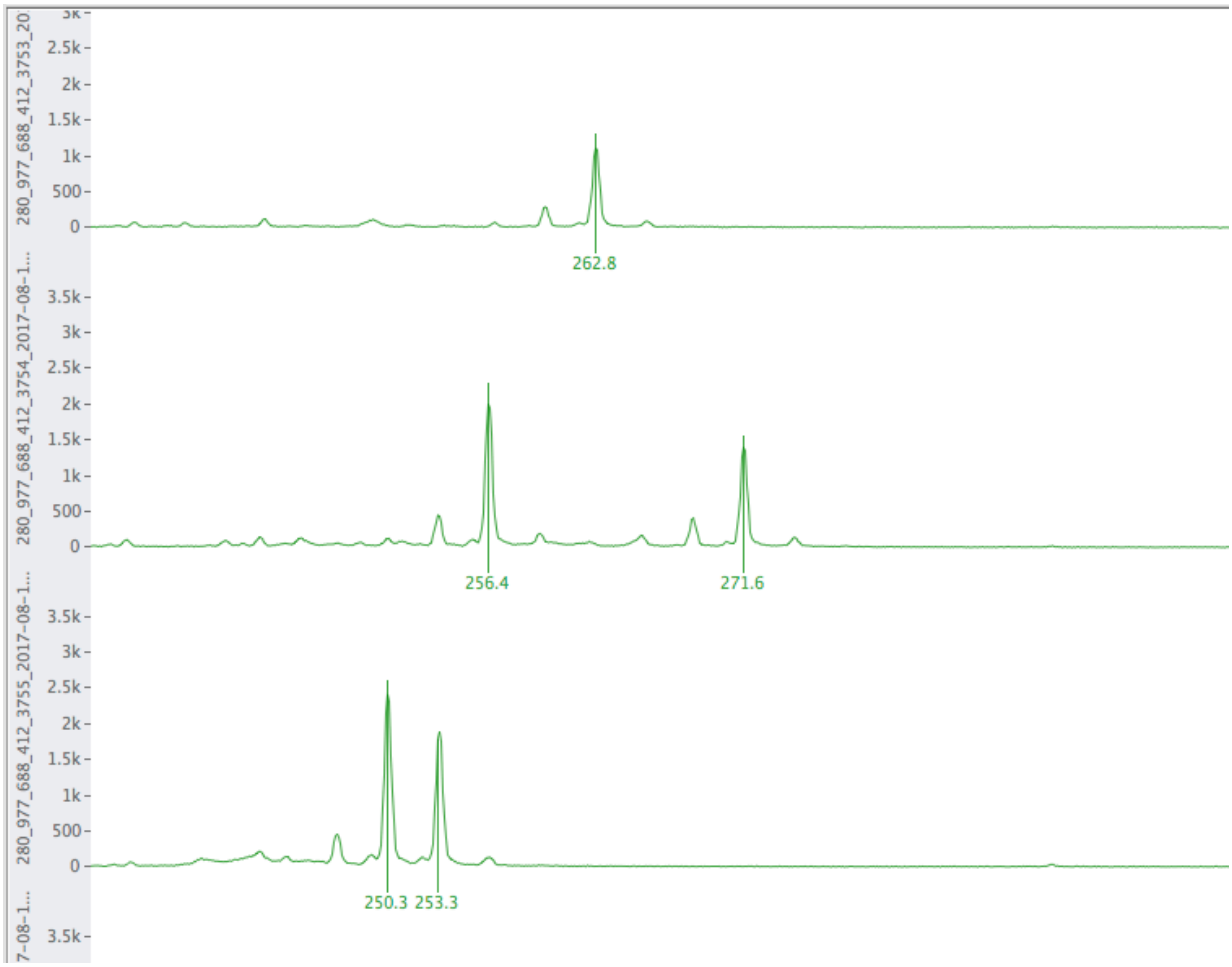
Sam	Site	LOH 931	LOH 931	LOH 1248	LOH 1248	LOH 936	LOH 936	LOH 684	LOH 684	LOH 412	LOH 412	LOH 663	LOH 663	LOH 749	LOH 749	LOH 868	LOH 868	LOH 1497	LOH 1497	LOH 688	LOH 688	LOH 505	LOH 505	LOH 865	LOH 865	LOH 1512	LOH 1512	M
3640	Tabayag	229	241	187	197	219	219	341	344	NA	NA	290	302	207	219	169	211	200	204	323	341	443	443	237	249	149	163	1
3644	Tabayag	241	241	197	203	219	219	341	341	247	247	304	310	221	231	151	151	208	208	320	320	440	443	243	243	163	163	0
3645	Tabayag	NA	NA	197	197	NA	NA	341	341	256	256	336	336	221	221	143	151	NA	NA	NA	NA	389	389	227	227	141	141	4
3646	Tabayag	NA	NA	197	203	219	219	341	341	247	247	304	310	209	221	151	151	208	208	NA	NA	440	443	243	243	163	163	2
3647	Tabayag	227	245	203	203	219	219	NA	NA	NA	NA	308	308	207	219	191	191	200	206	326	329	NA	NA	221	239	NA	NA	4
3648	Canbantug	243	243	197	199	219	219	341	341	274	274	288	310	209	221	167	175	200	204	320	320	NA	NA	243	243	171	171	1
3649	Canbantug	229	241	203	205	219	219	341	341	244	244	308	310	NA	NA	175	175	200	200	NA	NA	440	443	243	243	171	171	2
3650	Canbantug	229	241	205	205	219	219	341	341	244	244	290	308	247	247	175	197	200	208	NA	NA	NA	NA	243	243	167	171	2
3651	Canbantug	217	227	197	201	207	219	341	344	NA	NA	308	310	NA	NA	175	197	190	200	NA	NA	443	443	221	221	167	171	3
3652	Canbantug	227	229	201	201	219	219	341	344	247	247	290	308	221	241	157	207	202	204	320	329	437	440	243	245	163	171	0
3653	Canbantug	215	215	197	197	217	219	341	344	247	247	306	308	NA	NA	169	169	202	204	320	332	443	443	241	241	171	171	1
3654	Canbantug	NA	NA	197	199	219	219	341	344	262	262	288	290	207	217	NA	NA	200	202	320	320	437	440	241	241	171	171	2
3655	Canbantug	205	215	187	197	219	219	341	344	247	247	306	310	NA	NA	NA	NA	202	204	329	329	NA	NA	239	239	149	149	3
3656	Canbantug	205	215	199	201	219	219	344	344	244	247	290	306	NA	NA	151	159	198	208	320	320	NA	NA	243	243	171	171	2
3658	Canbantug	215	241	197	203	NA	NA	341	344	244	247	306	310	NA	NA	189	189	194	204	320	320	440	440	243	243	163	171	2
3659	Canbantug	225	243	191	201	209	219	341	341	247	247	290	308	209	243	155	155	192	202	320	326	NA	NA	237	253	NA	NA	2
3661	Canbantug	NA	NA	187	197	219	219	341	344	247	247	308	310	219	231	151	199	204	208	323	341	NA	NA	223	227	149	171	2
3664	Canbantug	NA	NA	197	201	219	219	341	344	NA	NA	290	310	237	237	199	199	200	204	332	335	443	443	227	229	171	171	2
3665	Canbantug	227	227	197	201	219	219	NA	NA	NA	NA	290	306	229	247	181	209	NA	NA	332	341	428	428	227	229	149	171	3
3670	Mag-alambac	225	225	191	201	207	219	344	344	247	247	306	306	209	243	157	161	204	208	320	329	437	437	243	243	167	175	0
3671	Mag-alambac	227	235	201	201	219	219	344	344	247	247	304	304	229	229	161	165	202	204	320	329	428	449	241	241	147	169	0
3672	Mag-alambac	225	225	201	201	219	219	344	344	247	247	304	304	221	221	165	181	200	200	320	338	428	467	239	239	151	167	0
3674	Mag-alambac	225	225	191	201	219	219	344	344	NA	NA	290	306	209	209	181	181	200	204	NA	NA	NA	NA	NA	NA	151	151	4
3676	Mag-alambac	227	227	197	201	219	219	341	341	NA	NA	288	288	221	221	159	181	200	202	320	329	440	440	239	241	147	171	1
3677	Mahangin	227	239	197	201	219	219	341	344	244	283	288	306	221	221	165	165	202	204	320	320	428	464	239	239	171	175	0
3678	Mahangin	215	215	201	201	207	207	344	344	244	244	304	304	243	247	183	183	202	204	320	329	449	449	239	239	171	171	0

Sam	Site	LOH 931	LOH 931	LOH 1248	LOH 1248	LOH 936	LOH 936	LOH 684	LOH 684	LOH 412	LOH 412	LOH 663	LOH 663	LOH 749	LOH 749	LOH 868	LOH 868	LOH 1497	LOH 1497	LOH 688	LOH 688	LOH 505	LOH 505	LOH 865	LOH 865	LOH 1512	LOH 1512	M
3681	Mahangin	227	227	197	197	219	233	341	344	247	247	288	304	231	243	161	165	200	200	320	320	NA	NA	NA	NA	171	171	2
3682	Mahangin	215	215	201	201	219	219	341	344	247	253	304	304	235	235	183	199	204	204	320	329	428	428	229	235	167	171	0
3685	Mahangin	227	255	201	201	219	219	341	344	247	253	288	302	261	263	169	201	202	204	329	341	443	443	221	221	173	175	0
3687	Cangcabalong	215	253	201	201	219	219	341	344	244	283	290	306	217	221	NA	NA	200	208	NA	NA	NA	NA	241	241	171	171	3
3691	Cangcabalong	225	225	197	197	197	219	341	344	NA	NA	304	304	221	241	185	185	200	208	308	320	NA	NA	241	241	171	171	2
3693	Cangcabalong	NA	NA	NA	NA	NA	NA	341	341	244	244	290	290	217	231	157	157	200	200	338	341	437	437	241	241	171	171	3
3499	Cangcabalong	225	227	201	201	219	219	344	344	NA	NA	304	308	221	259	155	155	NA	NA	308	326	449	449	241	243	167	171	2
3484	Mantalongon	225	225	191	201	219	219	341	344	247	247	288	310	221	221	183	205	NA	NA	320	323	440	440	229	229	173	175	1
3486	Mantalongon	215	226	191	201	219	219	344	344	NA	NA	288	312	NA	NA	165	165	200	202	320	329	437	437	239	239	173	173	2
3487	Mantalongon	227	241	201	201	219	219	341	344	253	253	NA	NA	221	221	193	195	NA	NA	320	323	467	467	245	245	171	173	2
3707	Baybayun	227	243	197	201	223	233	341	344	244	244	306	306	221	243	NA	NA	200	200	320	329	437	449	245	245	149	175	1
3710	Baybayun	NA	NA	NA	NA	NA	NA	344	344	247	247	304	306	261	261	157	157	NA	NA	320	320	449	449	229	249	169	171	4
3712	Baybayun	225	255	197	197	219	219	344	344	NA	NA	290	304	235	235	159	161	202	204	329	329	437	437	241	241	147	171	1
3713	Baybayun	225	225	197	201	207	219	341	341	253	253	306	306	221	239	NA	NA	202	204	329	329	434	434	227	243	167	171	1
3714	Baybayun	217	227	201	201	219	219	341	341	NA	NA	306	306	229	229	161	181	200	204	320	329	407	428	239	239	NA	NA	2
3715	Baybayun	217	227	201	201	219	219	344	344	247	247	NA	NA	219	231	159	191	NA	NA	320	329	NA	NA	243	243	167	167	3
3718	Baybayun	227	255	197	201	NA	NA	344	344	NA	NA	306	306	255	257	157	161	200	202	320	341	449	449	241	241	163	173	2
3721	Baybayun	227	253	197	197	219	219	341	344	229	235	306	306	221	241	161	161	202	202	320	329	NA	NA	241	241	163	171	1
3726	Upper Becerril	241	243	201	201	219	219	341	344	274	289	288	302	231	231	163	163	NA	NA	320	329	NA	NA	229	229	149	149	2
3727	Upper Becerril	215	243	193	203	NA	NA	344	347	250	289	290	302	231	231	157	181	202	202	329	329	434	434	NA	NA	149	149	2
3729	Upper Becerril	243	243	197	201	NA	NA	341	341	244	289	290	318	233	255	163	163	202	208	320	344	NA	NA	231	241	171	173	2
3750	Nangka	215	215	191	201	219	219	341	344	256	262	288	308	259	261	151	151	202	204	323	323	440	440	NA	NA	149	173	1
3752	San Antonio	243	243	193	203	NA	NA	341	344	265	283	NA	NA	233	233	195	195	202	202	320	338	440	440	NA	NA	147	173	3
3753	San Antonio	243	245	201	203	219	219	341	347	262	262	322	324	231	231	153	159	202	208	323	329	NA	NA	231	231	147	173	1
3754	San Antonio	231	231	191	201	NA	NA	341	344	256	271	288	290	233	261	195	195	202	202	320	323	440	440	231	231	147	167	1
3755	San Antonio	215	227	191	201	207	219	341	344	250	253	NA	NA	233	233	NA	NA	200	202	323	338	443	443	223	229	147	171	2

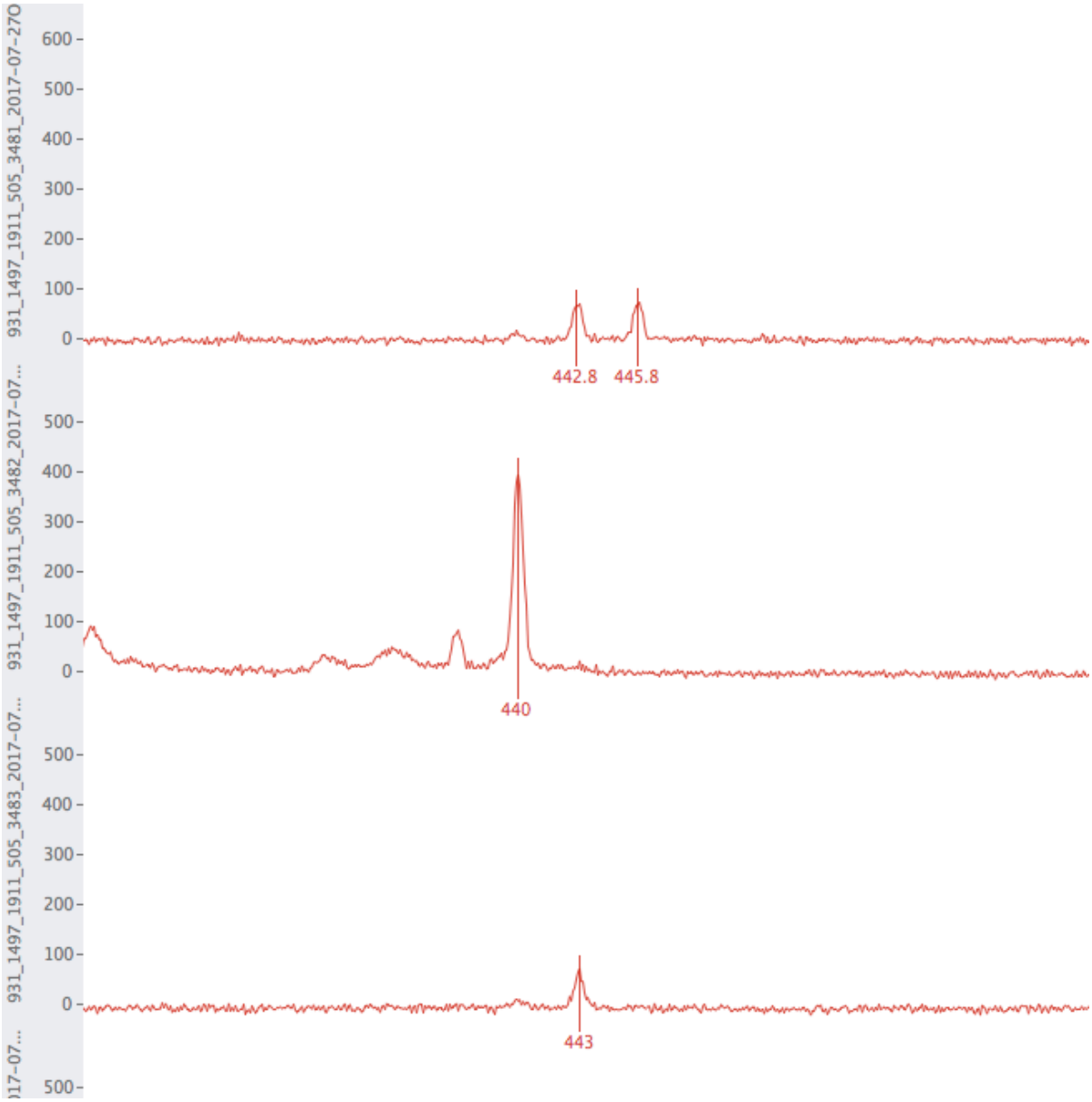
Sam	Site	LOH 931	LOH 931	LOH 1248	LOH 1248	LOH 936	LOH 936	LOH 684	LOH 684	LOH 412	LOH 412	LOH 663	LOH 663	LOH 749	LOH 749	LOH 868	LOH 868	LOH 1497	LOH 1497	LOH 688	LOH 688	LOH 505	LOH 505	LOH 865	LOH 865	LOH 1512	LOH 1512	M
3756	San Antonio	225	225	191	201	219	219	341	344	274	274	290	290	231	253	NA	NA	200	202	323	323	443	443	241	241	169	169	1
3757	San Antonio	215	243	191	201	219	219	341	344	271	271	288	288	255	275	NA	NA	NA	NA	323	323	440	440	NA	NA	149	149	3
3769	San Antonio	227	227	191	201	219	219	341	341	244	244	314	314	255	255	NA	NA	202	202	323	341	440	440	NA	NA	149	173	2
3774	Nug-as	243	243	191	201	219	219	341	341	271	271	300	300	233	233	159	159	200	200	320	326	NA	NA	229	241	151	167	1
3778	Nug-as	NA	NA	199	203	207	219	341	341	244	280	288	288	237	237	177	179	204	208	326	335	464	464	223	223	147	147	1
3779	Nug-as	243	243	197	201	NA	NA	341	341	NA	NA	NA	NA	NA	NA	155	181	200	208	320	338	464	464	229	239	147	147	4
3785	Nug-as	NA	NA	191	201	NA	NA	341	341	274	274	290	316	251	277	189	189	202	204	323	326	440	440	231	231	151	151	2
3787	Nug-as	227	249	191	201	207	219	341	344	259	259	NA	NA	265	267	163	163	202	208	338	338	NA	NA	231	241	NA	NA	3
3790	Nug-as	225	227	193	203	219	219	341	344	253	274	300	300	251	267	179	181	202	204	320	350	440	440	235	235	171	171	0
3798	Nug-as	215	227	193	203	219	219	341	344	256	256	300	318	235	251	163	163	200	202	320	323	463	463	229	247	147	173	0
3807	Nug-as	221	243	193	203	219	219	341	341	NA	NA	288	300	257	259	193	193	202	202	338	350	NA	NA	NA	NA	147	147	3
3817	Nug-as	243	243	191	201	207	219	344	344	247	274	290	308	255	257	159	159	202	202	323	323	464	464	NA	NA	151	171	1
3836	Nug-as	225	225	199	203	219	219	341	341	259	271	302	304	269	271	NA	NA	200	204	320	329	440	440	231	247	147	147	1
3843	Nug-as	243	243	193	203	NA	NA	341	341	256	259	288	288	219	219	185	187	202	202	323	323	461	461	231	245	149	149	1
3847	Nug-as	225	243	193	203	219	219	341	344	262	274	322	324	NA	NA	155	155	202	204	323	344	440	440	NA	NA	169	171	2
3849	Nug-as	245	245	199	203	207	219	341	341	259	259	288	288	231	231	181	191	200	200	320	320	464	464	NA	NA	171	173	1
3850	Nug-as	225	227	189	199	219	219	341	341	259	259	288	288	255	261	NA	NA	202	204	320	329	NA	NA	NA	NA	149	149	3
3851	Nug-as	243	243	193	203	207	219	341	344	274	274	304	304	245	255	167	167	NA	NA	320	320	NA	NA	NA	NA	NA	NA	4
3856	Poblacion	241	245	201	201	219	219	341	344	247	274	288	288	219	219	181	183	200	204	329	338	464	464	NA	NA	147	147	1
		LOH 931		LOH 1248		LOH 936		LOH 684		LOH 412		LOH 663		LOH 749		LOH 868		LOH 1497		LOH 688		LOH 505		LOH 865		LOH 1512		
	% Missing data	15.17		2.07		8.97		4.83		15.86		11.03		21.38		14.48		15.17		7.59		28.28		18.62		16.55		
	M13 primer	NED		6FAM		VIC		6FAM		VIC		VIC		VIC		NED		VIC		NED		PET		PET		PET		

Appendix 14. The morphology of select alleles of the 13 loci from a few individuals of Tetrastigma loheri*. Chromatograms were viewed in Geneious 6.1.8. The type of M13 fluorescent dye is indicated in parenthesis.*

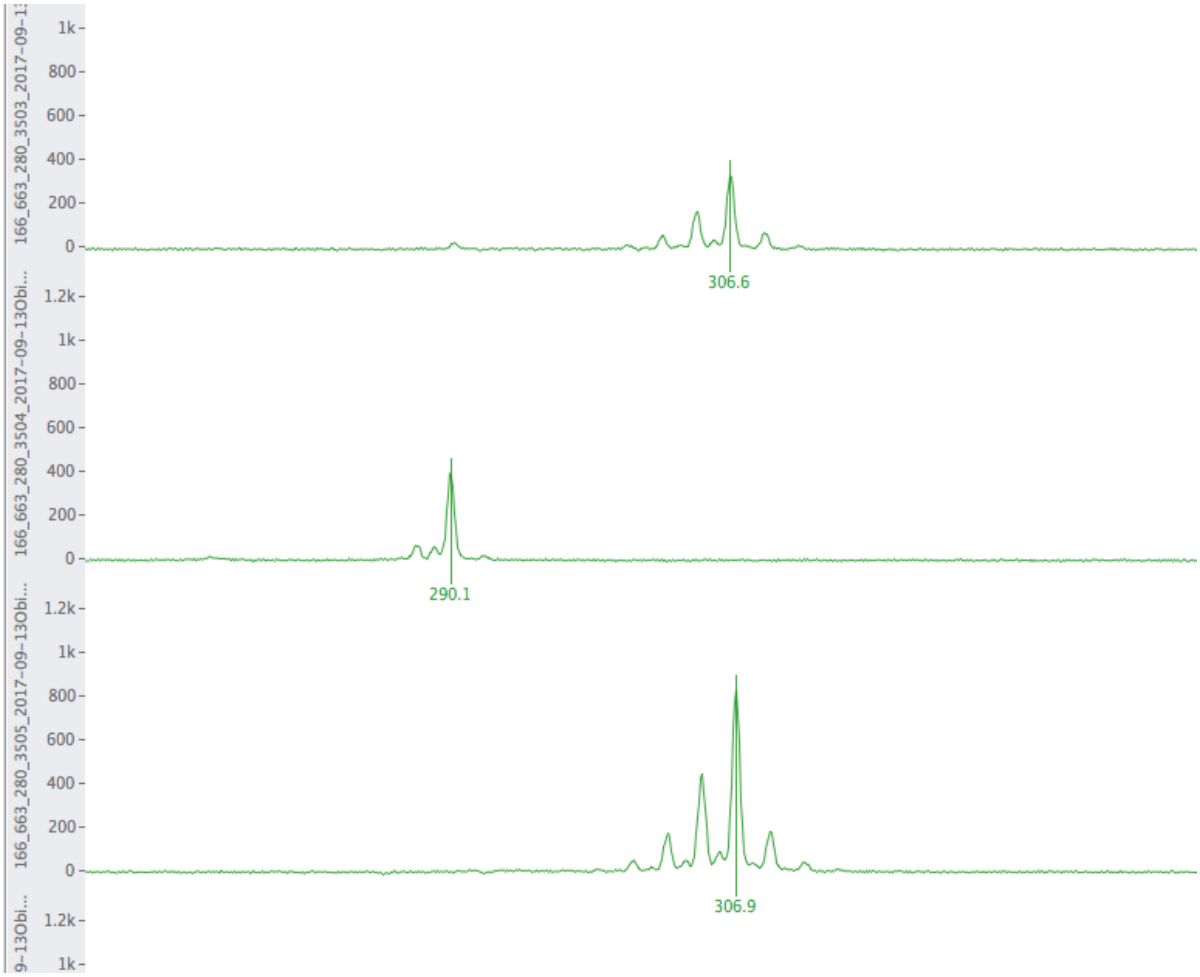
A. LOH 412 (VIC) for sample number 3753, 3754, and 3755.



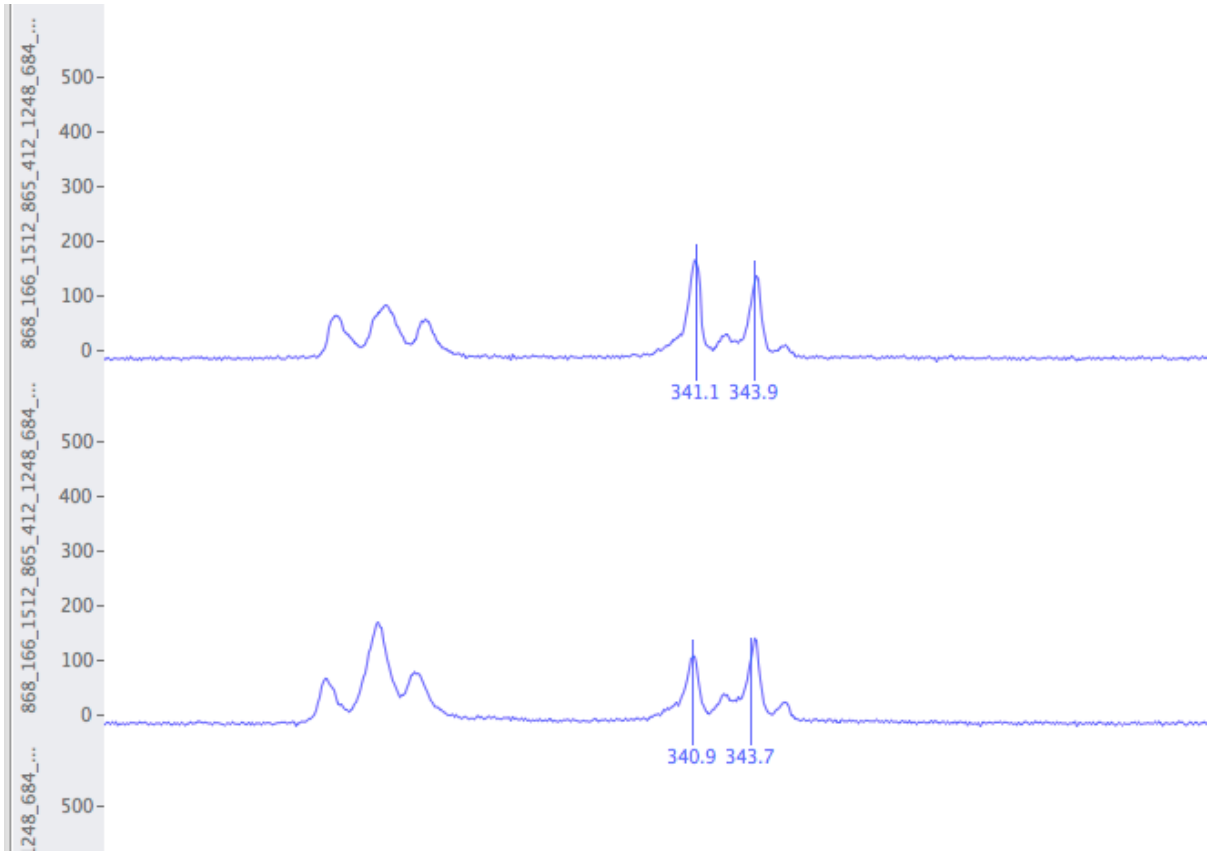
B. LOH 505 (PET) for sample number 3481, 3482, and 3483.



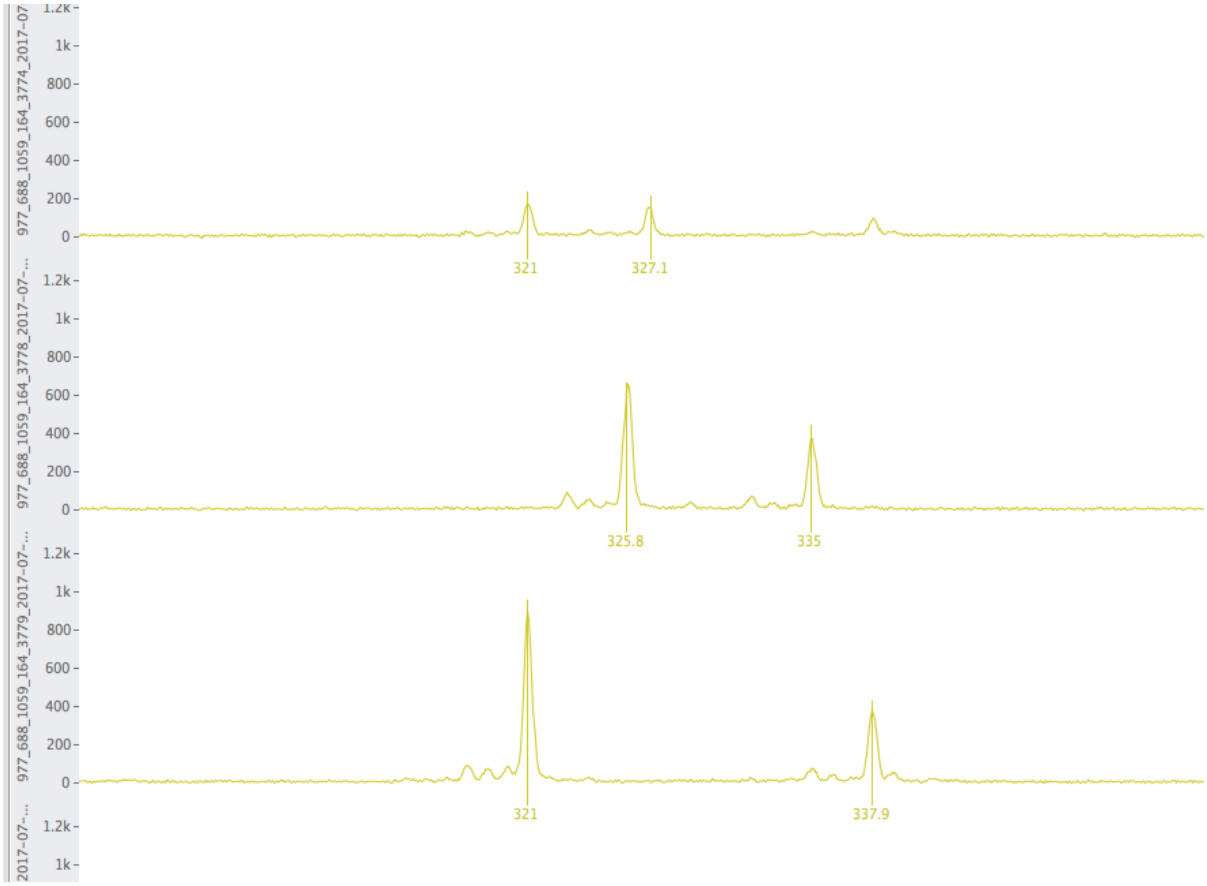
C. LOH 663 (VIC) for sample number 3503, 3504, and 3505.



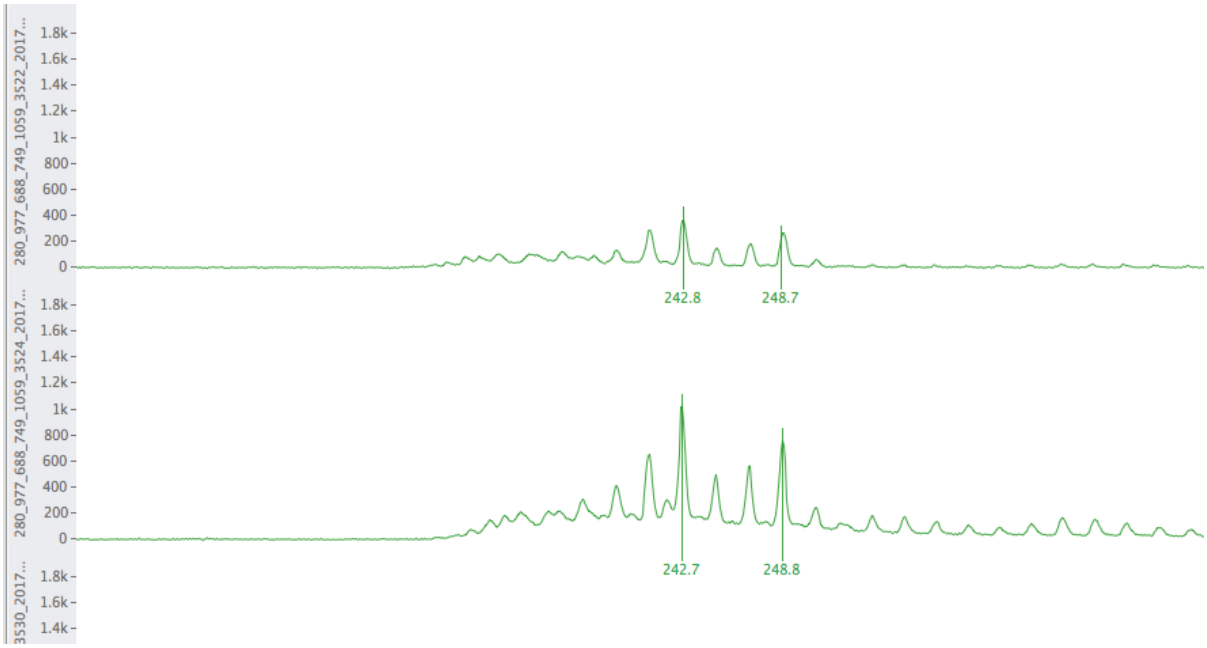
D. LOH 684 (6FAM) for sample number 3530 and 3531.



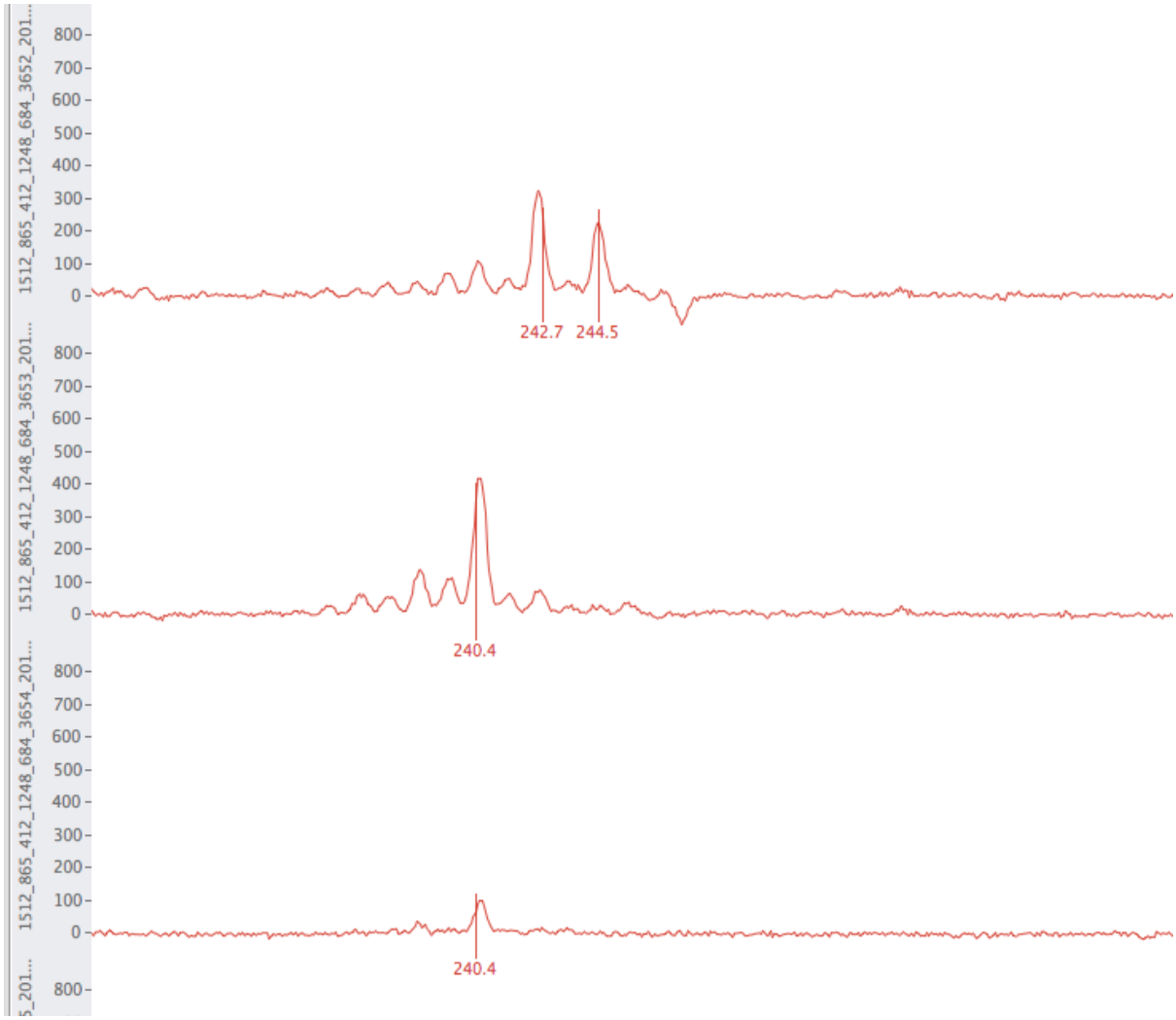
E. LOH 688 (NED) for sample number 3774, 3778, and 3779.



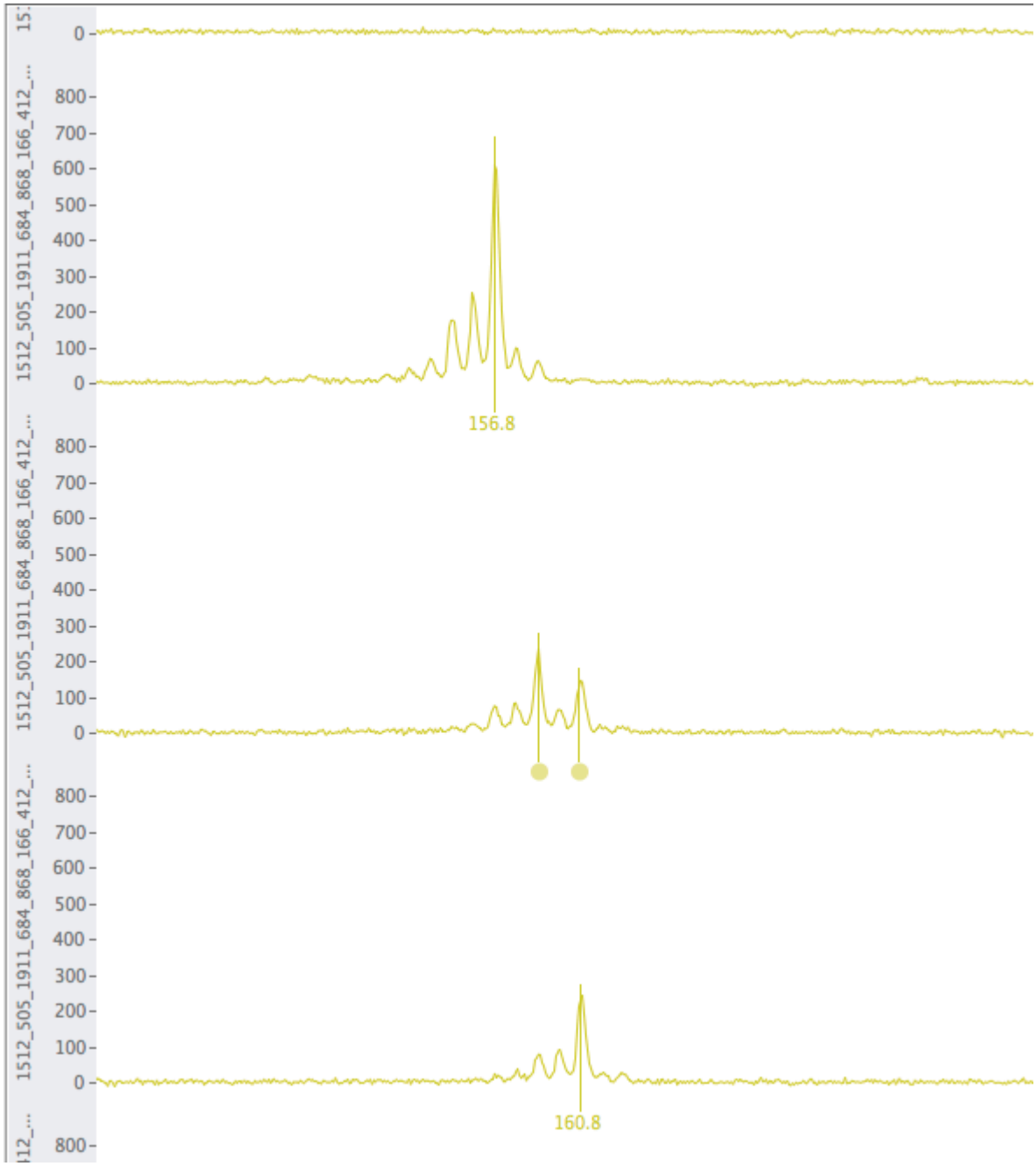
F. LOH 749 (VIC) for sample number 3522 and 3524.



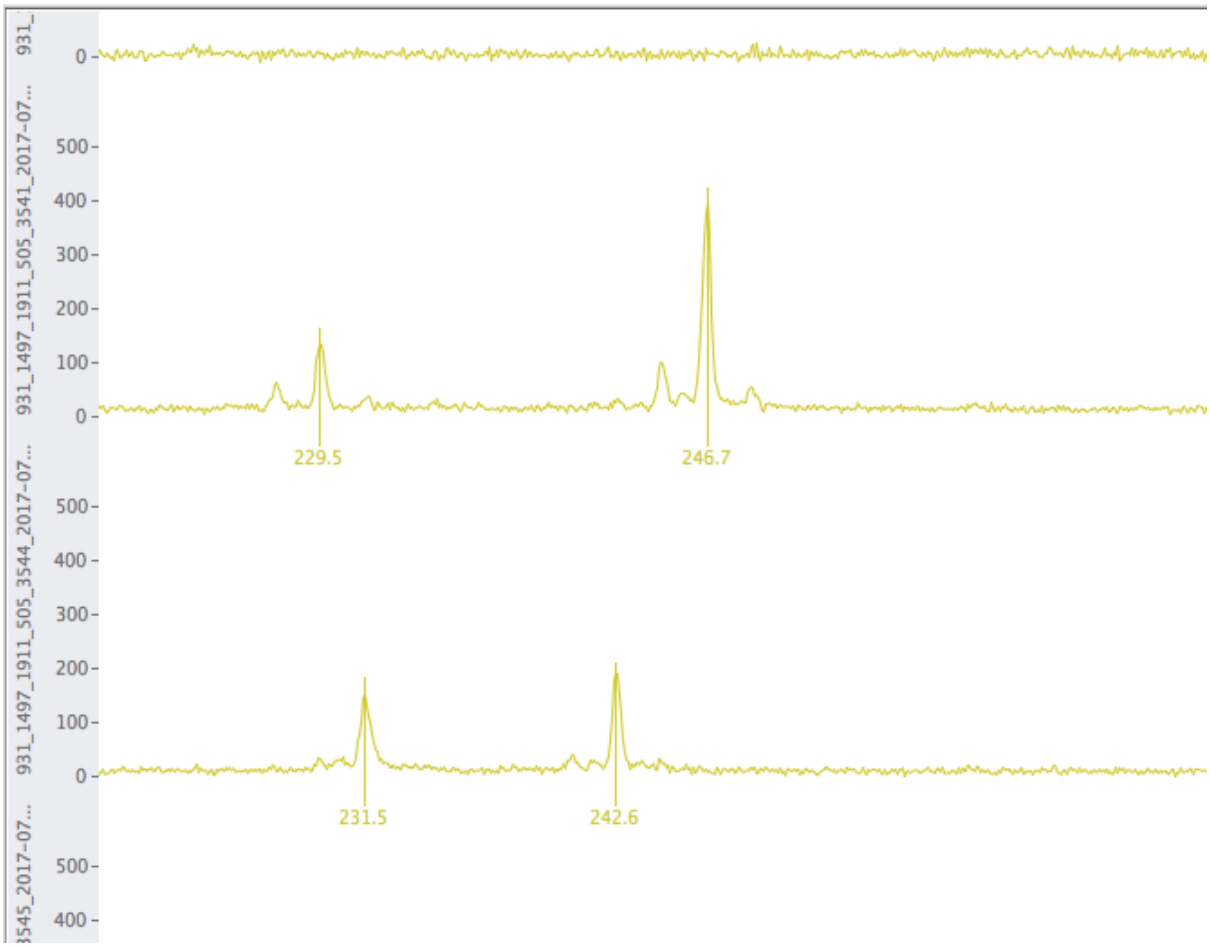
G. LOH 865 (PET) for sample number 3652, 3653, and 3654.



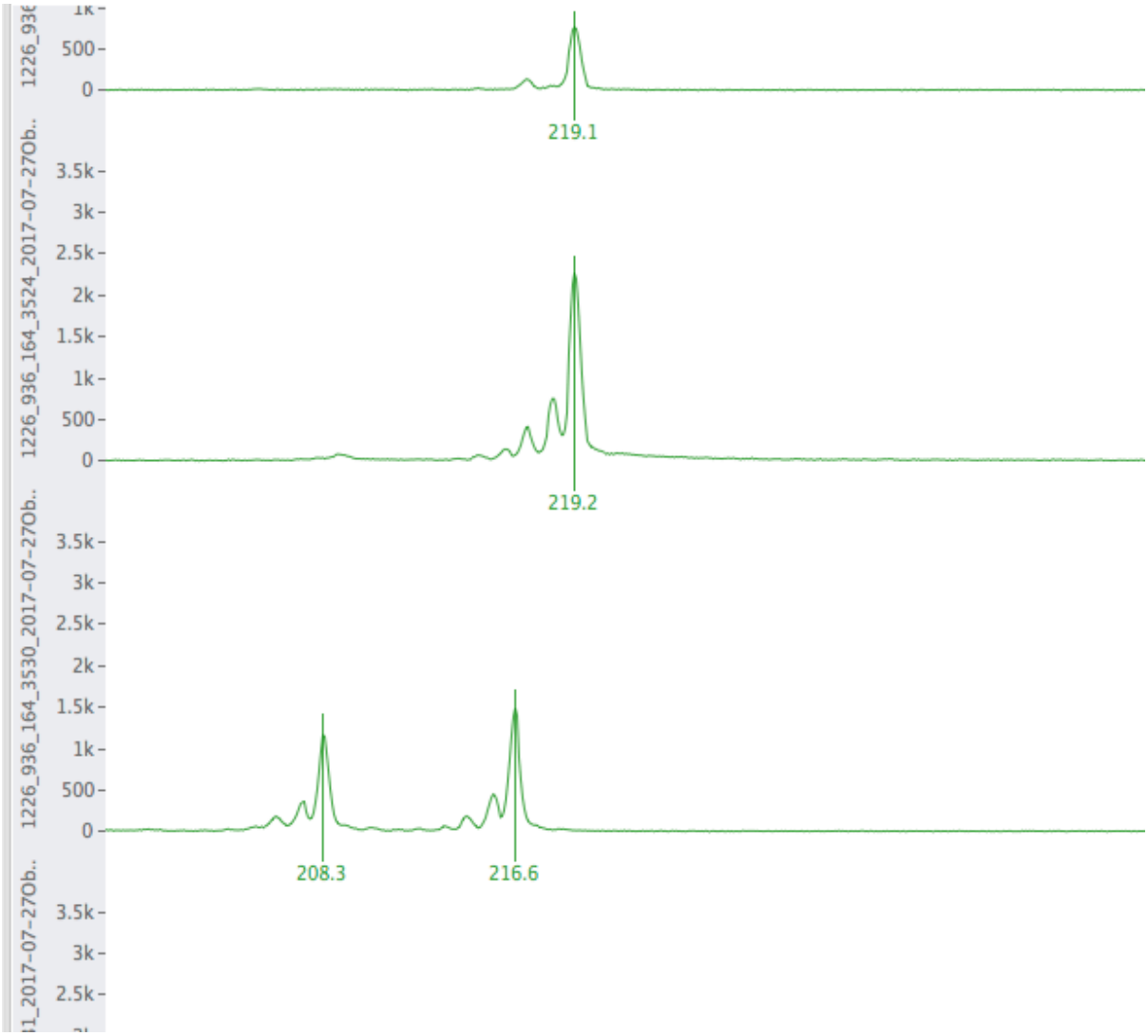
H. LOH 868 (NED) for sample number 3710, 3712, and 3721.



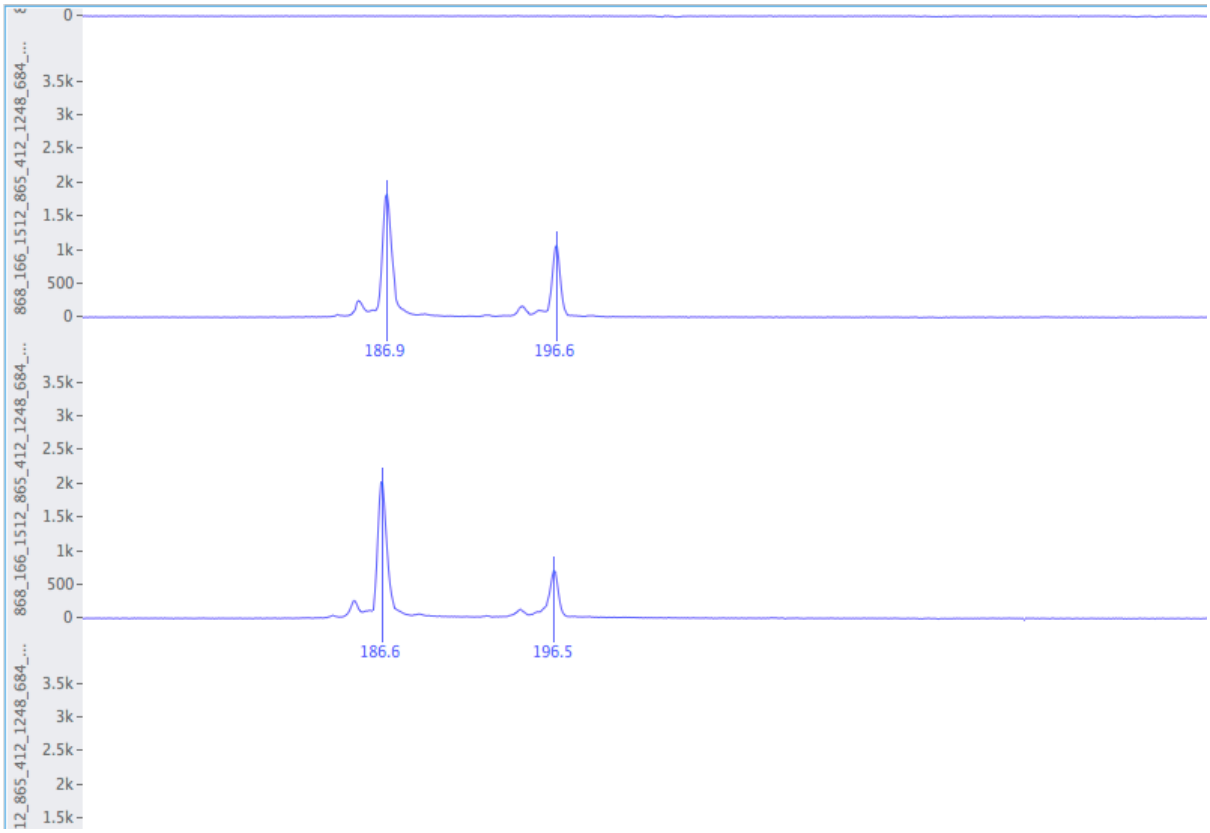
I. LOH 931 (NED) for sample number 3541 and 3544.



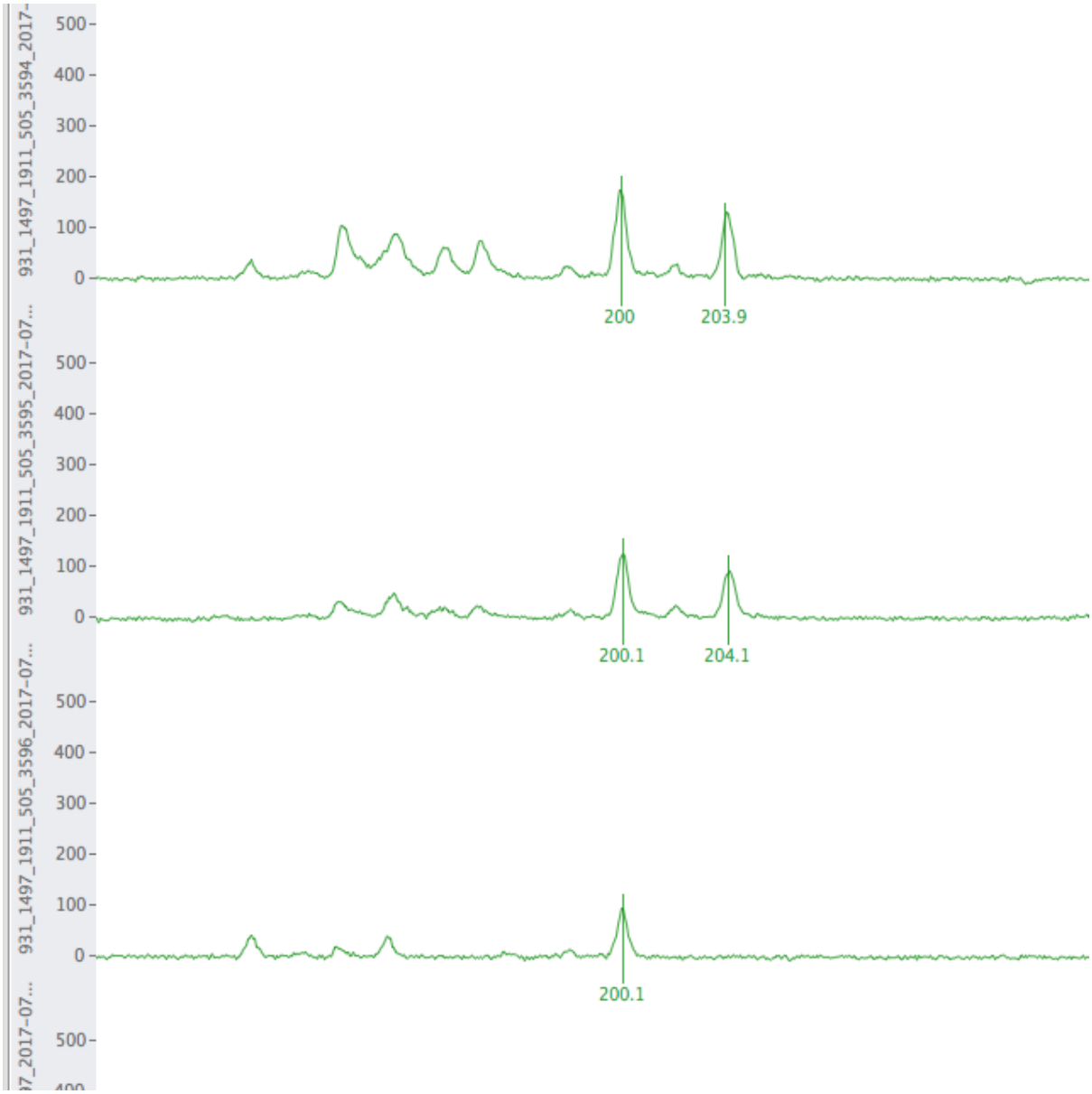
J. LOH 936 (VIC) for sample number 3524 and 3530.



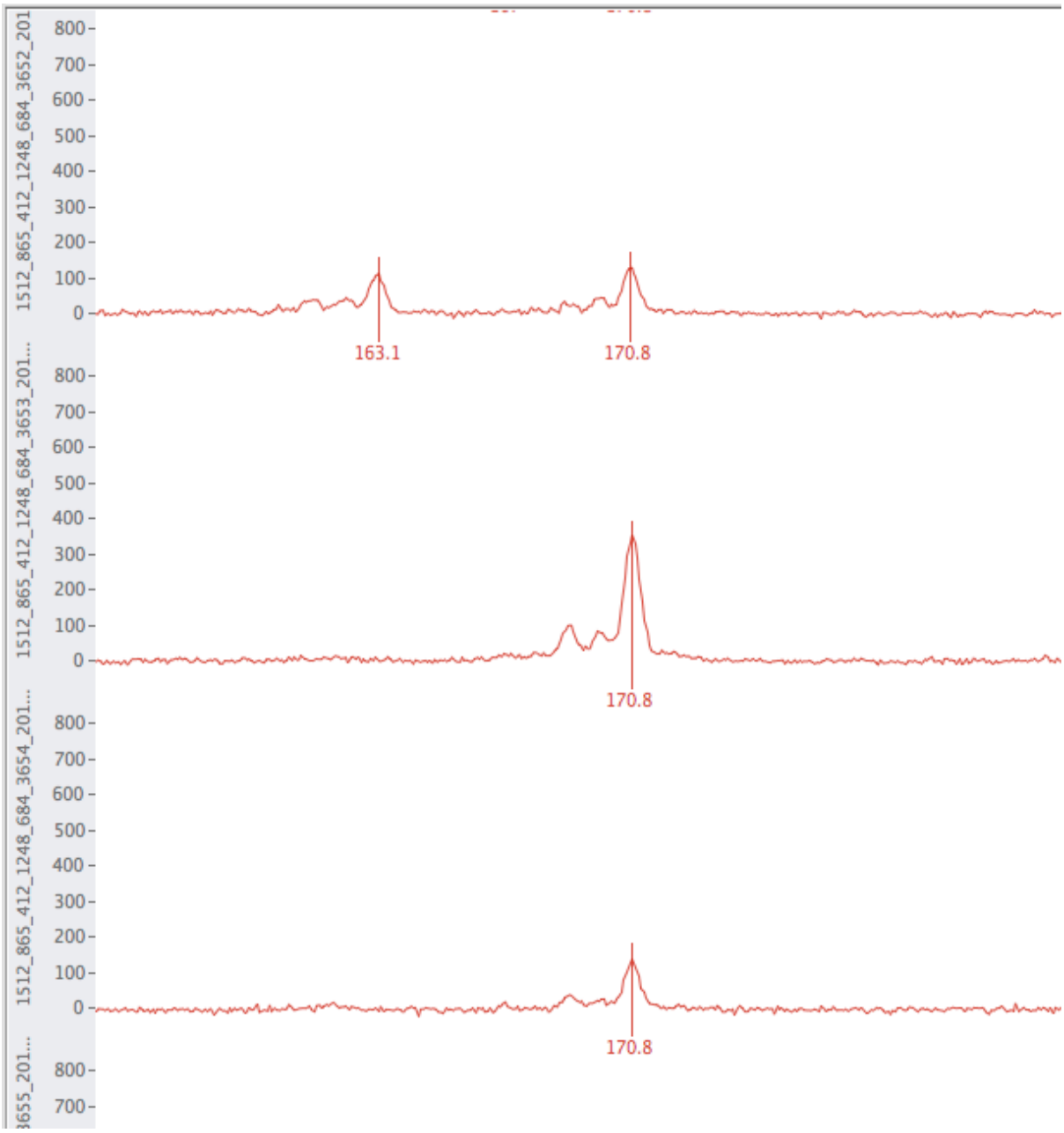
K. LOH 1248 (6FAM) for sample number 3530 and 3531.



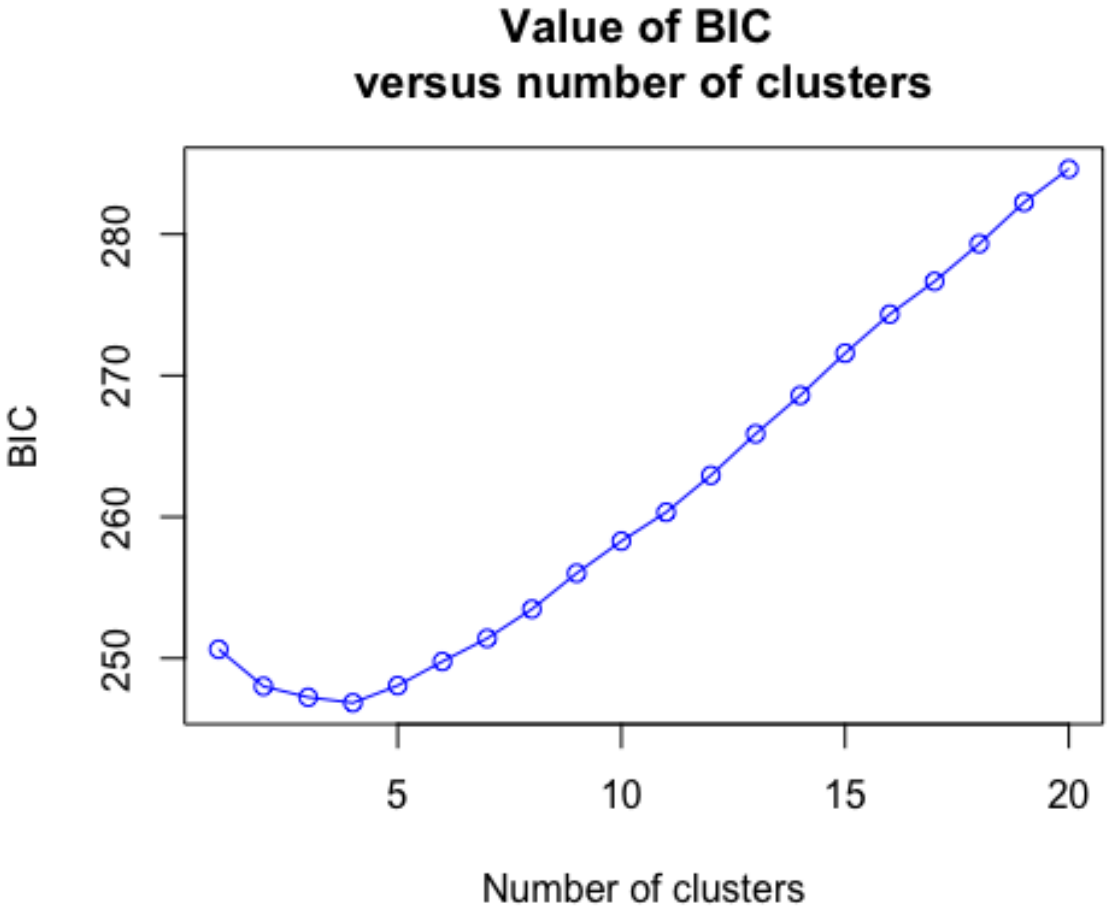
L. LOH 1497 (VIC) for sample number 3594, 3595, and 3596.



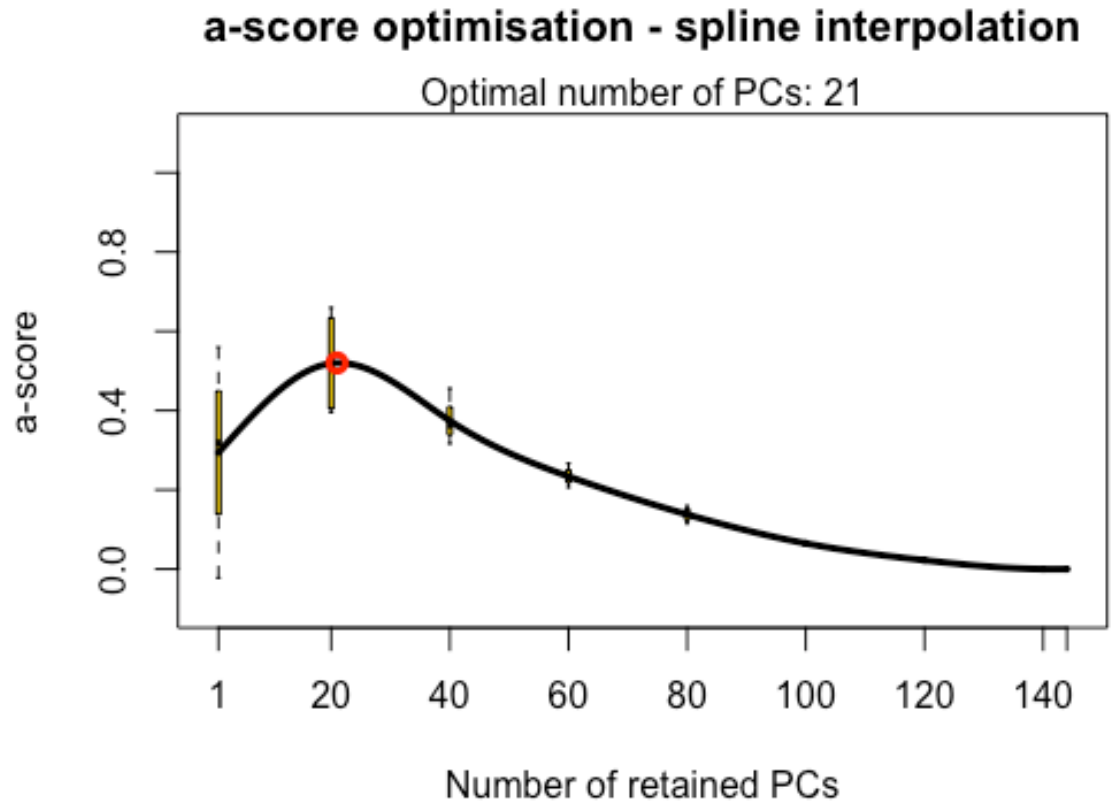
M. LOH 1512 (PET) for sample number 3652, 3653, and 3654.



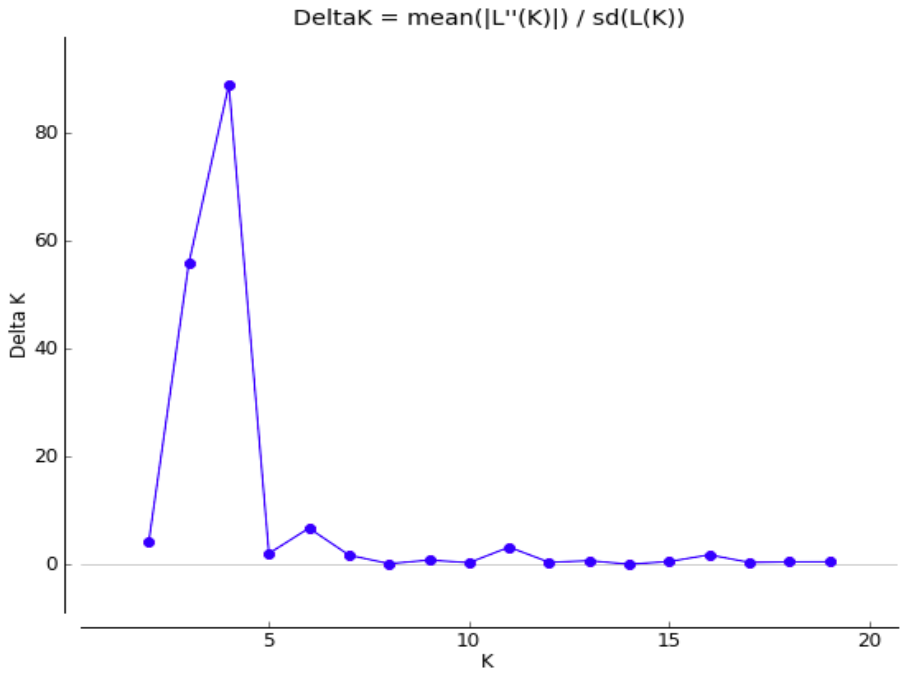
Appendix 15. Plot of Bayesian Information Criterion values and the number of clusters. The inferred best number of clusters is indicated by the lowest BIC value which is 4.



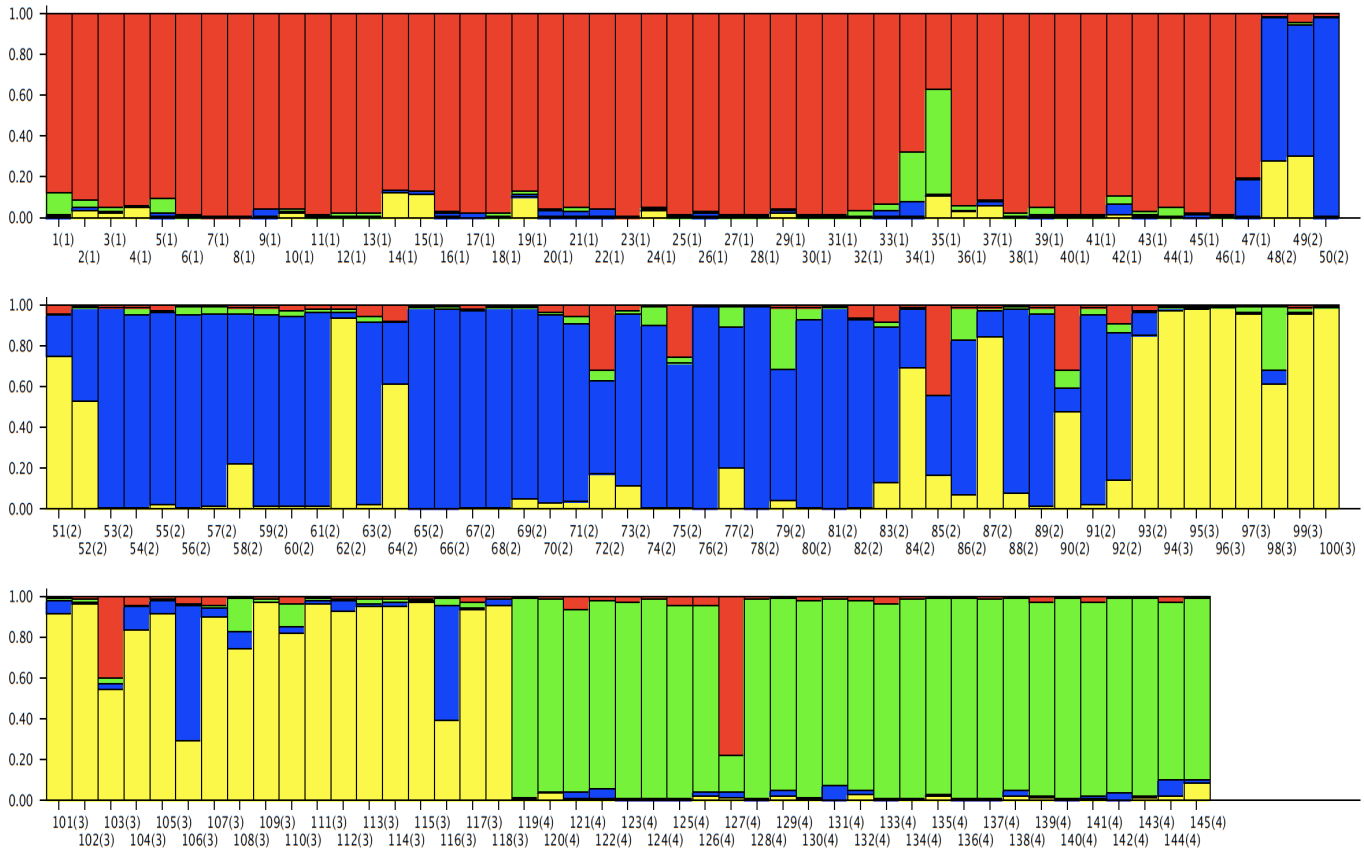
Appendix 16. The optimal number of principal components (PCs) to be retained in the DAPC analysis using a-score values as indicated by the red circle.



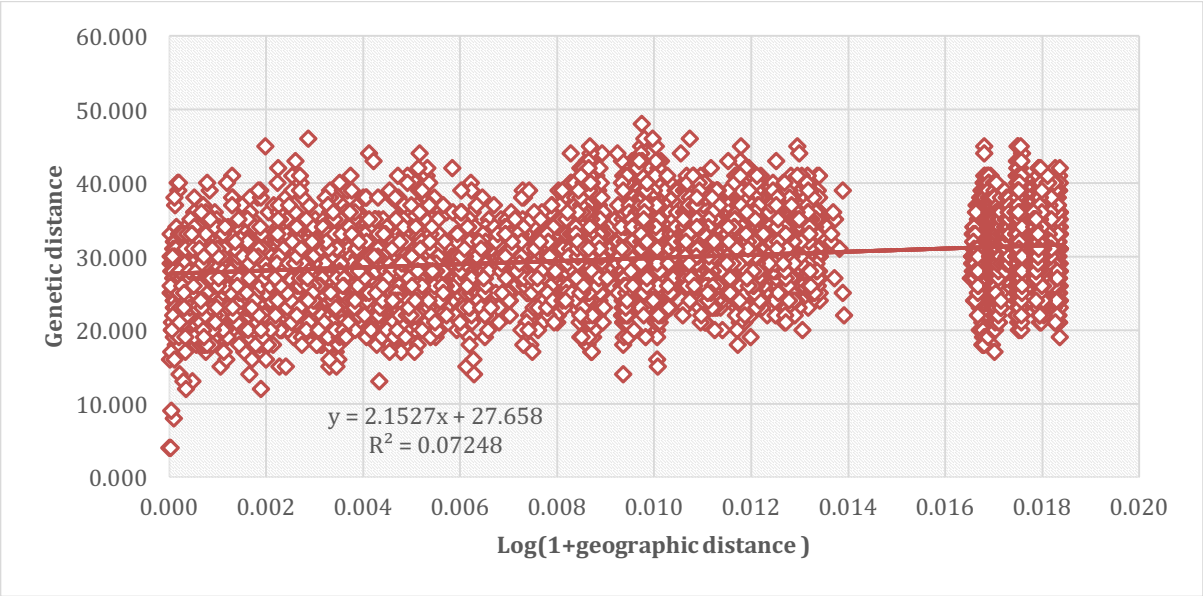
Appendix 17. Inferred optimal number of k based on the highest value of delta k using the Evanno et al. method (2005).



Appendix 18. Detailed bar plot of membership probabilities of Tetrastigma loheri *individuals generated by CLUMPAK -Clustering Markov Packager Across K on the web using results from STRUCTURE. Horizontal axis shows the number assigned to individuals based on their order in the datasheet. Identified migrants at 0.1 migration prior from Table 4.6 were as follows with their corresponding sample number in parenthesis: individual 51 (Obico 409), individual 62 (Obico 427), individual 90 (Obico 458), individual 103(Obico 484). The number in the parenthesis refers to the sampling origin of the individuals where 1=CCPL, 2=Argao, 3=Dalaguete, and 4=Alcoy. Vertical axis shows the membership probability values.*



Appendix 19. Mantel correlation between pairwise genetic distances of Tetrastigma loheri *individuals and log (1+ geographic distances). The R value of 0.269 is significant (p=0.001).*



Appendix 20. A forest landscape in the Central Cebu Protected Landscape showing the proximity of a village to the forest.

Appendix 21. A forest in Dalaguete next to the road and with signs of land clearing.

Appendix 22. Aerial view of the typical landscape intervening the four areas. This area is found between Dalaguete and Argao. Photo taken from Google Earth. Accessed on 31 May 2019.